vol. 3   no. 3   2015

MATHEMATICS AND MECHANICS
of
Complex Systems

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS

msp.org/memocs

STATIONARY SOLUTIONS OF KELLER–SEGEL-TYPE CROWD MOTION AND HERDING MODELS: MULTIPLICITY AND DYNAMICAL STABILITY

JEAN DOLBEAULT, GASPARD JANKOWIAK AND PETER MARKOWICH

In this paper we study two models for crowd motion and herding. Each of the models is of Keller–Segel type and involves two parabolic equations, one for the evolution of the density and one for the evolution of a mean field potential. We classify all radial stationary solutions, prove multiplicity results, and establish some qualitative properties of these solutions, which are characterized as critical points of an energy functional. A notion of variational stability is associated with such solutions.

Dynamical stability in the neighborhood of a stationary solution is also studied in terms of the spectral properties of the linearized evolution operator. For one of the two models, we exhibit a Lyapunov functional which allows us to make the link between the two notions of stability. Even in that case, for certain values of the mass parameter, with all other parameters taken in an appropriate range, we find that two dynamically stable stationary solutions exist. We further discuss the qualitative properties of the solutions using theoretical methods and numerical computations.

1. Introduction

The Keller–Segel model in chemotaxis has attracted lots of attention over recent years. However, most of the theoretical results have been obtained either in a parabolic-elliptic setting or when the coefficients, such as the chemosensitivity coefficient, are independent of the solution. Models used in biology usually involve coefficients which depend on the solution itself, thus making the problems far more nonlinear, and also far less understood. The crowd motion and herding models considered here are two problems in the same class, where the main additional features, compared to the standard version of the Keller–Segel model, are the limitation (the prevention of overcrowding) of drift for the mass density in both models, and the limitation of the source in the equation for the chemoattractant

in one of the two models. Such limitations have important consequences: there are multiple solutions for a given mass, in certain regimes; plateau-like solutions have an interesting pattern for modeling issues; and the flux limitation forbids concentration and guarantees nice solution properties, but also raises nontrivial stability issues concerning the set of stationary solutions, which we investigate numerically. The two models can be considered as test cases for the understanding of a very large class of parabolic-parabolic systems with the property of having several attractors. The fact that radial solutions are bounded and can be fully parametrized in relatively simple terms makes the study tractable. Most of the difficulties come from the complicated dependence of the solutions on the total mass, which is the crucial parameter in the two cases. Numerically, the difficulty comes from the parameters of the model, which have to be chosen in ranges that make the problem rather stiff.

1.1. *Description of the models.* We consider herding and crowd motion models describing the evolution of a density ρ of individuals subject to a drift ∇D and confined to a bounded, open set $\Omega \subset \mathbb{R}^d$. The evolution equation for ρ is given by

$$\partial_t \rho = \Delta \rho - \nabla \cdot (\rho(1-\rho)\nabla D), \tag{1}$$

where ρ_t stands for the derivative of ρ with respect to time t and $\rho(1-\rho)$ includes the term for the prevention of overcrowding. For an isolated system, it makes sense to introduce a no-flux boundary condition, that is,

$$(\nabla \rho - \rho(1-\rho)\nabla D) \cdot \nu = 0 \quad \text{on} \quad \partial\Omega, \tag{2}$$

which guarantees the conservation of the number of individuals (or conservation of mass), namely that

$$\int_\Omega \rho \, dx = M \tag{3}$$

is independent of t. In the models considered in this paper, we shall assume that the potential D solves a parabolic equation

$$\partial_t D = \kappa \Delta D - \delta D + g(\rho) \tag{4}$$

and is subject to homogeneous Neumann boundary conditions

$$\nabla D \cdot \nu = 0 \quad \text{on} \quad \partial\Omega. \tag{5}$$

We restrict our purpose either to *model (I)*, when

$$g(\rho) = \rho(1-\rho), \tag{6}$$

or to *model (II)*, when

$$g(\rho) = \rho. \tag{7}$$

In this paper, our purpose is to characterize stationary solutions and determine their qualitative properties.

1.2. *Motivations.*

Human crowd motion models are motivated by the desire to prevent stampedes in public places, mainly by implementing better walkway design. Most crowd motion models do not convey herding effects well enough, that is, loosely speaking, when people bunch up and try to move in the same direction, as typically occurs in emergency situations.

In an effort to improve herding and crowd motion models, Burger et al. [2011] have derived models (I) and (II) as the continuous limits of a microscopic cellular automaton model introduced in [Kirchner and Schadschneider 2002]. This takes the form a parabolic-parabolic system for a density of people ρ and for field D, where D is a mean field potential which carries the herding effects. Basically, people are subject to random motion, with a preference for moving in the direction others are following. Random effects are taken into account by a diffusion, while a drift is created by the potential D, which accounts for locations that are or were previously occupied. To account for the packing of the people, empty spaces are preferred, which explains the role of the $(1 - \rho)$ term in front of the drift, with 1 being the maximal density. Such a correction is referred to as *prevention of overcrowding* in the mathematical literature.

Both quantities ρ and D undergo diffusion, which happens much faster for ρ, this point being reflected by the fact that the constant κ is assumed to be small. The potential D decays over time with rate $\delta > 0$ and increases proportionally to the density ρ, but only if the density is not too high in the case of model (I); this is taken into account by the source term $g(\rho)$ given either by (6) or (7). As we shall see, interesting phenomena also occur when δ is small.

In many aspects, these models are quite similar to the Keller–Segel model used in chemotaxis. The prevention of overcrowding has already been considered in several papers, either in the parabolic-elliptic case [Burger et al. 2008; 2010] or the parabolic-parabolic case [Di Francesco and Rosado 2008] (with diffusion-dominated large-time asymptotics) and [Burger et al. 2010] (where, additionally, the case of several species and cross-diffusion was taken in to account). In these papers the emphasis was put on asymptotic behaviors, with a discussion of the possible asymptotic states and behaviors depending on nonlinearities in [Burger et al. 2006] and a study of plateau-like quasistationary solutions and their motion in [Burger et al. 2008]. This of course makes sense when the domain is the entire space, but a classification of the stationary solutions in bounded domains and in particular plateau-like solutions is still needed, as it is strongly suggested by

[Burger et al. 2011] that such solutions have interesting properties, for instance, in terms of stability.

Because of the $(1 - \rho)$ factor in front of the drift, the transport term vanishes in our models as ρ approaches 1, so that for any initial data bounded by 1 the density remains bounded by 1. Hence blow-up, which is a major difficulty for the analysis of the usual Keller–Segel system for masses over 8π (see, for instance, [Blanchet et al. 2006]), does not occur here. In contrast with the parabolic-elliptic Keller–Segel model with prevention of overcrowding studied in [Burger et al. 2010], models (I) and (II) are based on a system of coupled parabolic equations. This has interesting consequences for the evolution problem as, for example, it introduces memory effects. It also has various consequences for the dynamical stability of the stationary states. In model (I), the source term in the equation for D involves $\rho(1 - \rho)$ instead of ρ. This nonlinear source term introduces additional difficulties: for instance, no Lyapunov functional is known.

1.3. *Main results.* Let us summarize some of the main results of this paper, in the cases of models (I) and (II), when Ω is a ball, as far as radial nonnegative stationary solutions are concerned. As we shall see below the stationary solutions of interest are either constants or monotone functions, which are then plateau-like.

Theorem 1. *Let Ω be a ball and consider solutions of models (I) and (II) subject to boundary conditions (2) and (5). Then the masses of the radial nonnegative stationary solutions as defined by (3) range between 0 and $|\Omega|$ and we have*:

(i) *Nonconstant stationary solutions exist only for M in a strict subinterval $(0, |\Omega|)$.*

(ii) *Constant solutions are variationally and dynamically unstable in a strictly smaller subinterval.*

(iii) *There is a range of masses in which only nonconstant stationary solutions are stable, given by the condition that $\kappa\lambda_1 + \delta$ is small enough, where λ_1 denotes the lowest positive eigenvalue of $-\Delta$ in Ω subject to Neumann homogeneous boundary conditions.*

(iv) *For any given mass, variationally stable stationary solutions with low energy are either monotone or constant; in the case of model (II), monotone, plateau-like solutions are then stable and attract all low-energy solutions of the evolution problem in a certain range of masses.*

Much more can be said on stationary solutions, as we shall see below, and some of our results are not restricted to radial solutions on a ball. The natural parameter for the solutions of models (I) and (II) is M, but it is much easier to parametrize the set of solutions by an associated Lagrange multiplier; see Section 2. In particular, stationary solutions are then critical points of an energy defined in Section 3, and there is a notion of *variational stability* associated with this energy. Taking into

account the mass constraint, as done in Section 4, makes the problem more difficult. To study the evolution problem, one can rely on a Lyapunov functional introduced in Section 5, but only in the case of model (II). *Dynamical stability* is studied through the spectrum of the linearized evolution operator in Section 6 and the interplay between notions of variational and dynamical stability is also studied in detail. How to harmonize the two points of view on stability is a question that models (I) and (II) share with all parabolic-parabolic models of chemotaxis. In the case of model (II), results are summarized in Theorem 26. The issue of the stability of monotone — constant or nonconstant — solutions is a subtle question and most of this paper is devoted to this point. Precise definitions of variational and dynamical stability will be given later on.

Numerical results go beyond what can be proved rigorously. Because we use the parametrization by the Lagrange multiplier, we are able to compute *all* radial solutions. In practice, we shall focus on the role of constant and monotone plateau-like solutions. A list of detailed qualitative results is provided at the beginning of Section 7. Theoretical and numerical results are discussed in Section 8.

1.4. *Some references.* The two models considered in this paper have been introduced in [Burger et al. 2011] at the partial differential equation level. Considerations on the stability of constant solutions can be found therein as well. Models (I) and (II) involve a system of two parabolic equations, like the so-called parabolic-parabolic Keller–Segel system, for which we primarily refer to [Calvez and Corrias 2008]. In such a model, stationary solutions have to be replaced by self-similar solutions, which also have multiplicity properties (see [Biler et al. 2011]). How the parabolic-parabolic model is related to the parabolic-elliptic case has been studied in [Biler and Brandolese 2009; Calvez and Corrias 2008]. The parabolic-elliptic counterpart of model (I) is known: for plateau solutions and the coarsening of the plateaus, we refer to [Burger et al. 2008] (also see [Burger et al. 2006; 2010]; related models can be found in the literature labeled as Keller–Segel models with logistic sensitivity or congestion models).

One of the technical but crucial issues for a complete classification of all solutions is how to parametrize the set of solutions. Because Lyapunov or energy functionals are not convex, this is a far more difficult issue than in the repulsive case, for which we refer to [Dolbeault et al. 2001]. The lack of convexity makes it difficult to justify but, at a formal level, the evolution equations in model (II) can be interpreted as gradient flows with respect to some metric involving a Wasserstein distance (see [Blanchet et al. 2015] in the case of the Keller–Segel model and [Blanchet and Laurençot 2013] for a more general setting; also see [Laurençot and Matioc 2013] for an earlier result in the same spirit). To be precise, one has to consider the Wasserstein distance for ρ and a L^2 distance for D as in [Calvez and

Carrillo 2012]. The difficulty comes from the fact that the Lyapunov functional is not displacement convex (see, for instance, [Blanchet et al. 2008] and subsequent papers in the parabolic-elliptic case of the Keller–Segel system). Using methods introduced in [Matthes et al. 2009] this may eventually be overcome, but it is still open at the moment, as far as we know.

2. Radial stationary solutions

2.1. *A parametrization of all radial stationary solutions.* Any stationary solution of (1) solves

$$\nabla\rho - \rho(1-\rho)\nabla D = 0 \quad \text{on} \quad \Omega,$$

which means

$$\rho = \frac{1}{1 + e^{-\phi}}, \tag{8}$$

where $\phi = D - \phi_0$ and $\phi_0 \in \mathbb{R}$ is an integration constant determined by the mass constraint (3); ϕ_0 is the unique real number such that

$$\int_\Omega \frac{1}{1 + e^{\phi_0 - D}} \, dx = M. \tag{9}$$

Taking into account boundary conditions (5), (4) now amounts to

$$-\kappa\Delta\phi + \delta(\phi + \phi_0) - f(\phi) = 0 \quad \text{on} \quad \Omega \tag{10}$$

with boundary conditions

$$\nabla\phi \cdot \nu = 0 \quad \text{on} \quad \partial\Omega. \tag{11}$$

The functions f and F are defined by $f = F'$ and

$$F(\phi) = \rho = \frac{1}{1 + e^{-\phi}} \quad \text{and} \quad f(\phi) = \rho(1-\rho) = \frac{e^{-\phi}}{(1 + e^{-\phi})^2} \quad \text{in the model (I) case,}$$

$$F(\phi) = \log(1 + e^{\phi}) \quad \text{and} \quad f(\phi) = \rho = \frac{1}{1 + e^{-\phi}} \quad \quad \text{in the model (II) case.}$$

The crucial observation for our numerical computation is based on the following result.

Proposition 2. *If Ω is the unit ball in $\mathbb{R}^d$, $d \geq 2$, all radial solutions of (10) and (11) with f as above are smooth and can be found by solving the shooting problem*

$$-\kappa\left(\varphi_a'' + \frac{d-1}{r}\varphi_a'\right) + \delta(\varphi_a + \phi_0) - f(\varphi_a) = 0, \quad \varphi_a'(0) = 0, \quad \varphi_a(0) = a,$$

as a function of the parameter $a \in \mathbb{R}$. The shooting criterion is $\varphi_a'(1) = 0$.

If $d = 1$, all solutions in $\Omega = (0, 1)$ are given by the above ordinary differential equation (ODE).

Proof. The proof presents no difficulty and is left to the reader. □

2.2. *Constant solutions.* Determining ϕ such that $\delta(\phi + \phi_0) - f(\phi) = 0$, that is,

$$k(\phi) := \frac{1}{\delta} f(\phi) - \phi = \phi_0, \tag{12}$$

exactly amounts to determining the (possibly multivalued) function $\phi_0 \mapsto k^{-1}(\phi_0)$. The following result is not restricted to the special case of f as defined in model (I) or (II).

Lemma 3. *Let $\delta > 0$. Assume that $f \in C^1(\mathbb{R})$ is bounded and $\lim_{\phi \to \pm\infty} f'(\phi) = 0$. Then the function $\phi \mapsto k'(\phi) = (1/\delta) f'(\phi) - 1$ has 2ℓ zeros for some $\ell \in \mathbb{N}$ and* (12) *has at most $2\ell + 1$ solutions. Moreover, for $|\phi_0|$ large enough,* (12) *has one and only one solution, which is such that ρ given by* (8) *converges to 0 as $\phi_0 \to +\infty$ and to 1 as $\phi_0 \to -\infty$.*

If $|\phi_0|$ is large, we observe that $k(\phi) \sim -\phi$. Other properties are elementary consequences of the intermediate values theorem and are left to the reader. A plot is shown in Figure 1.

With $f = F'$ and f corresponding either to model (I) or (II), all assumptions of Lemma 3 are satisfied with $\ell = 0$ or 1. For use later, let us define

$$\phi_-(\phi_0) := \min k^{-1}(\phi_0) \quad \text{and} \quad \phi_+(\phi_0) := \max k^{-1}(\phi_0)$$
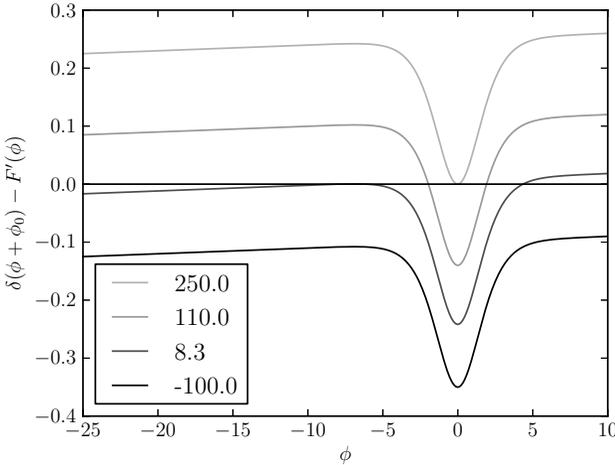
Figure 1. Plot of $\phi \mapsto \delta(\phi + \phi_0) - f(\phi) = \delta(\phi + \phi_0) - F'(\phi) = \delta(\phi_0 - k(\phi))$ for various values of ϕ_0. Each zero of the function provides a constant stationary solution of (1)–(5). The plot shown here corresponds to model (I), with $\delta = 10^{-3}$.

and emphasize that $\phi_\pm$ depend on ϕ_0. The set $k^{-1}(\phi_0)$ is reduced to a point if and only if $\phi_-(\phi_0) = \phi_+(\phi_0)$. From Lemma 3, we also know that

$$\phi_0^- := \inf\{\phi_0 \in \mathbb{R} : \phi_-(\phi_0) < \phi_+(\phi_0)\} \quad \text{and} \quad \phi_0^+ := \sup\{\phi_0 \in \mathbb{R} : \phi_-(\phi_0) < \phi_+(\phi_0)\}$$

are both finite.

Instead of parametrizing solutions by ϕ_0, it is interesting to think in terms of mass. Here is a first result (see Figure 2) in this direction, which follows from the property that $k'(\phi_\pm(\phi_0)) < 0$ for any $\phi_0 \in \mathbb{R}$.

Lemma 4. *Under the assumptions of Lemma 3, $\phi_0 \mapsto \phi_\pm(\phi_0)$ is monotone decreasing, and the corresponding masses are also monotone decreasing as functions of ϕ_0.*

The proof is elementary and left to the reader. If ϕ is a constant solution, it is a monotone increasing function of the mass according to (8). Hence the mass of a constant extremal solution $\phi = \phi_\pm(\phi_0)$ is a monotone decreasing function of ϕ_0. Moreover, we have

$$f'(\phi) = \rho(1 - \rho)h(\rho),$$
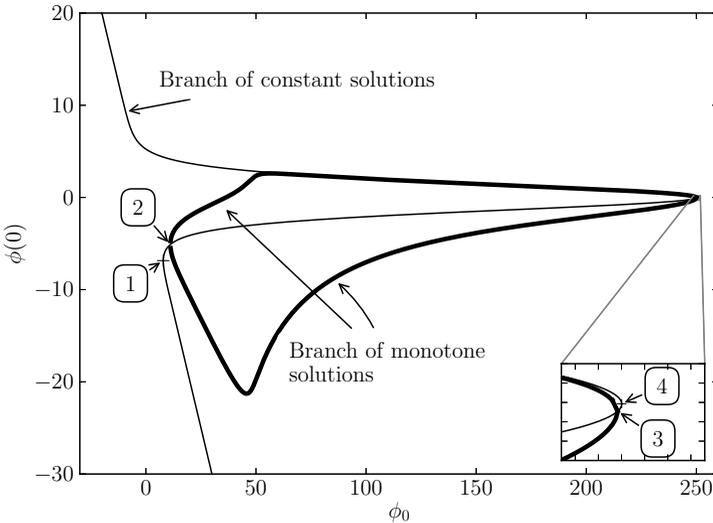
Figure 2. Parametrization by ϕ_0 of the branches of solutions in the case of model (I), $d = 1$, with $\delta = 10^{-3}$, $\kappa = 5 \times 10^{-4}$, and $\Omega = (0, 1)$. There are either one or three constant solutions for a given value of ϕ_0. Strictly monotone solutions correspond to the bold curve. Notice that on the upper part of the graph the two branches are close but distinct.

with ρ given by (8), $h(\rho) = 1 - 2\rho$ in the case of model (I), and $h(\rho) = 1$ in the case of model (II). A simple computation shows that $\mathsf{m} := \max_{\rho \in [0,1]} \rho(1-\rho)h(\rho)$ is equal to $1/(6\sqrt{3})$ and $1/4$ in the cases of models (I) and (II), respectively. As a consequence, with the notation of Lemma 3, $\ell = 0$ if either $\delta \geq \mathsf{m}$ or $\delta < \mathsf{m}$ and $\phi_0 \in \mathbb{R} \setminus (\phi_0^-, \phi_0^+)$. If $\delta < \mathsf{m}$ we find that $\ell = 1$ if $\phi_0 \in (\phi_0^-, \phi_0^+)$; there are exactly three constant solutions.

In the case of models (I) and (II), the (unique) constant solution taking values in $(\phi_-(\phi_0), \phi_+(\phi_0))$ is monotone increasing as a function of ϕ_0 (when it exists), thus defining a range of masses in which Theorem 1(iii) holds, as we shall see below.

3. Unconstrained energy and constant solutions

In this section we consider the problem for fixed ϕ_0. On the space $\mathrm{H}^1(\Omega)$, let us define the *energy functional* by

$$\mathscr{E}_{\phi_0}[\phi] := \frac{\kappa}{2} \int_\Omega |\nabla \phi|^2 \, \mathrm{d}x + \frac{\delta}{2} \int_\Omega |\phi + \phi_0|^2 \, \mathrm{d}x - \int_\Omega F(\phi) \, \mathrm{d}x. \tag{13}$$

It is clear from (3) that stationary solutions of models (I) and (II) are critical points of $\mathscr{E}_{\phi_0}$ (see Lemma 5) for some given Lagrange multiplier ϕ_0. Moreover, for a given ϕ_0, we know how to compute all radial solutions as explained in Section 2. Hence we shall first fix ϕ_0, study the symmetry of the minimizers of $\mathscr{E}_{\phi_0}$, and clarify the role of constant solutions.

3.1. *Critical points.*

Lemma 5. *Assume that F is Lipschitz continuous and Ω is bounded with $C^{1,\alpha}$ boundary for some $\alpha > 0$. With ϕ_0 kept constant, ϕ is a solution of (10) and (11) if and only if it is a critical point of $\mathscr{E}_{\phi_0}$ in $\mathrm{H}^1(\Omega)$.*

It is straightforward to check that $\mathscr{E}_{\phi_0}$ has a minimizer for any given ϕ, but such a minimizer is actually constant as we shall see in Corollary 8. Nonconstant solutions are therefore not minimizers of $\mathscr{E}_{\phi_0}$, for fixed ϕ_0. The regularity of the solution of (10) and (11) depends on the regularity of F, but when it is smooth as in the case of models (I) and (II), the standard elliptic theory applies and ϕ is smooth up to the boundary. We refer, for instance, to [Brezis 2011] as a standard reference book. Details are left to the reader and we shall assume without further notice that solutions are smooth from now on.

Notice that our original problem is not set with ϕ_0 fixed, but with mass constraint (3). Understanding how results for a given ϕ_0 can be recast into problems with M fixed is a major source of difficulties and will be studied in particular in Section 4.

3.2. *Linearized energy functional.* Consider the linearized energy functional

$$\lim_{\varepsilon \to 0} \frac{\mathscr{E}_{\phi_0}[\phi + \varepsilon \psi] - \mathscr{E}_{\phi_0}[\phi]}{2\varepsilon^2} = \int_{\Omega} \psi \, (\mathrm{E}_{\phi} \psi) \, \mathrm{d}x,$$

where ϕ is a stationary solution, $\psi \in \mathrm{H}^2(\Omega)$, and $\mathrm{E}_{\phi} \psi := -\kappa \Delta \psi + \delta \psi - F''(\phi) \psi$. Notice that with ρ given by (8), we have

$$\mathrm{E}_{\phi} \psi = -\kappa \Delta \psi + \delta \psi - \rho(1 - \rho) h(\rho) \psi, \tag{14}$$

with $h(\rho) = 1 - 2\rho$ in the case of model (I) and $h(\rho) = 1$ in the case of model (II).

3.3. *Stability and instability of constant solutions.* Denote by $(\lambda_n)_{n \in \mathbb{N}}$ the sequence of all eigenvalues of $-\Delta$ with homogeneous Neumann boundary conditions, counted with multiplicity. The eigenspace corresponding to $\lambda_0 = 0$ is generated by the constants. Three constant solutions coexist when constant solutions ϕ take their values in $k \circ (k')^{-1}(0, +\infty)$, that is, when

$$\delta - \rho(1 - \rho) h(\rho) < 0.$$

A constant solution $(\rho, D = \phi + \phi_0)$ is *variationally* unstable if E_{ϕ} has a negative eigenvalue, that is, if

$$\kappa \lambda_1 + \delta - \rho(1 - \rho) h(\rho) < 0. \tag{15}$$

When such a condition is satisfied, the constant solution ϕ cannot be a local minimizer of $\mathscr{E}_{\phi_0}$. *Dynamical* stability of the constant solutions with respect to the evolution governed by (1)–(5) will be studied in Section 6; in the case of constant solutions, such an instability is also determined by (15), as we shall see in Proposition 18.

Condition (15) is never satisfied if $\kappa \lambda_1 + \delta \geq \mathsf{m} := \max_{\rho \in [0,1]} \rho(1 - \rho) h(\rho)$. Otherwise, this condition determines a strict subinterval of $(0, 1)$ in terms of ρ, and hence an interval in ϕ. This proves Theorem 1(ii). A slightly more precise statement goes as follows.

Lemma 6. *Let $\delta > 0$. The set of values of ϕ_0 for which there are constant solutions of* (10) *which satisfy* (15) *with ρ given by* (8) *is contained in (ϕ_0^-, ϕ_0^+). Moreover, if there exists a constant, variationally unstable solution, then there is also a constant, variationally stable solution of* (10) *for the same value of ϕ_0, but with lower energy.*

The proof of Lemma 6 requires some additional observations. It will be completed in Section 3.6.

3.4. *Numerical range.* Cases of numerical interest studied in this paper are the following.

(1) In dimension $d = 1$ with $\Omega = (0, 1)$, the first unstable mode is generated by $x \mapsto \cos(\pi x)$ and corresponds to $\lambda_1 = \pi^2 \approx 9.87$.

(2) In dimension $d = 2$, the first positive critical point of the first Bessel function of the first kind J_0, that is, $r_0 := \min\{r > 0 : J_0'(r) = 0\}$, is such that $r_0 \approx 3.83$ so that $\lambda_{0,1} = r_0^2 \approx 14.68$ is an eigenvalue associated with the eigenspace generated by $r \mapsto J_0(rr_0)$. Applied to (15), this determines the range of *radial variational instability*. Recall that J_0 is the solution of $J_0'' + (1/r)J_0' + J_0 = 0$.

   Notice that nonradial instability actually occurs in a larger range, since the first positive critical point of the second Bessel function of the first kind J_1, that is, $r_1 := \min\{r > 0 : J_1'(r) = 0\}$, is such that $r_1 \approx 1.84$ so that $\lambda_{1,0} = r_1^2 \approx 3.39$ is an eigenvalue associated with the eigenspace generated by $r \mapsto J_1(rr_1)$, and $\lambda_1 = \lambda_{1,0} < \lambda_{0,1}$. Applied to (15), this determines the range of *variational instability*. Recall that J_1 is the solution of $J_1'' + (1/r)J_1' - (1/r^2)J_1' + J_1 = 0$.

The values of $\max_{\rho \in [0,1]} \rho(1 - \rho)h(\rho)$ are in practice also rather small, namely $1/(6\sqrt{3}) \approx 0.096$ and $1/4 = 0.25$ in the cases of models (I) and (II), respectively, which in practice, in view of the values of λ_1, makes the numerical computations rather stiff. In this paper we are interested in the qualitative behavior of the solutions and the role of the dimension, but not so much in the role of the surrounding geometry; hence we shall restrict our study to radial solutions. One of the advantages of dealing only with radial solutions is that we can use accurate numerical packages for solving ODEs and rely on shooting methods, thus getting a precise description of the solution set. Taking into account the effects of the geometry is another challenge but is, in our opinion, secondary compared to establishing all qualitative properties that can be inferred from our numerical computations. Another reason for restricting our study to radially symmetric functions is Proposition 2: using the shooting method, we have the guarantee of the description of all solutions, with additional information like the knowledge of the range in which to adjust the shooting parameter, as a consequence of the observations of Section 2.2 (see also Proposition 9). Within the framework of radial solutions, we can henceforth give a thorough description of the set of solutions, which is clearly out of reach in more general geometries. However, inasmuch as we deal with theoretical results, we will not assume any special symmetry of the solutions unless necessary.

In practice, the numerical computations of this paper are done with $\delta = 10^{-3}$ and κ ranging from 5×10^{-4} to 10^{-2}. Such small values are dictated by (15). They are also compatible with the computations and modeling considerations found in [Burger et al. 2011]. See Figure 2 for a plot corresponding to a rather generic diagram representing constant solutions for model (I) in dimension $d = 1$. Numerically, our interest lies in the nonconstant radial solutions that bifurcate from the constant solutions ϕ at threshold values for condition (15), that is, for values of ϕ_0

such that $\kappa\lambda_1 + \delta - \rho(1-\rho)h(\rho) = 0$, with $\lambda_1 = \pi^2$, in dimension $d = 1$, and $\lambda_1 = \lambda_{0,1}$ when $d = 2$. We shall take ϕ_0 as the bifurcation parameter and compute the mass of the solution only afterwards, thus arriving at a simple parametrization of all solutions. Our main results are therefore a complete description of branches of solutions bifurcating from constant ones and giving rise to *plateau* solutions. See Figure 3 for some plots of the solutions. We notice that in the range considered for the parameters, the transition from high to low values is not too sharp. The numerical study will be confined to radial monotone solutions, but we will briefly explain in Section 4.3 (at least when $d = 1$) what can be expected for a solution with several plateaus. Concerning stability issues, decomposition on appropriate basis sets will be required, as will be explained in Section 7.

3.5. *Qualitative properties of the stationary solutions.*

Lemma 7. *Let Ω be a bounded open set in $\mathbb{R}^d$ with C^2 boundary and assume that $k : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous with*

$$\liminf_{u \to -\infty} k(u) > 0 \quad and \quad \limsup_{u \to +\infty} k(u) < 0.$$

Assume that all zeros of k are isolated and denote them by $u_1 < u_2 < \cdots < u_N$ for some $N \geq 1$. Then any solution of class C^2 of $\Delta u + k(u) = 0$ in Ω satisfying $\nabla u \cdot x = 0$ on $\partial\Omega$ takes values in $[u_1, u_N]$.

Proof. Let $x^* \in \overline{\Omega}$ be a maximum point of u. We know that $-\Delta u(x^*) = k(u(x^*)) \geq 0$, even if $x^* \in \partial\Omega$, because of the boundary conditions. By assumption, we find that $u(x) \leq u(x^*) \leq u_N$ for any $x \in \overline{\Omega}$. Similarly, one can prove that $u \geq u_1$. $\square$

Applying Lemma 7 to (10) and (11) has straightforward but interesting consequences.
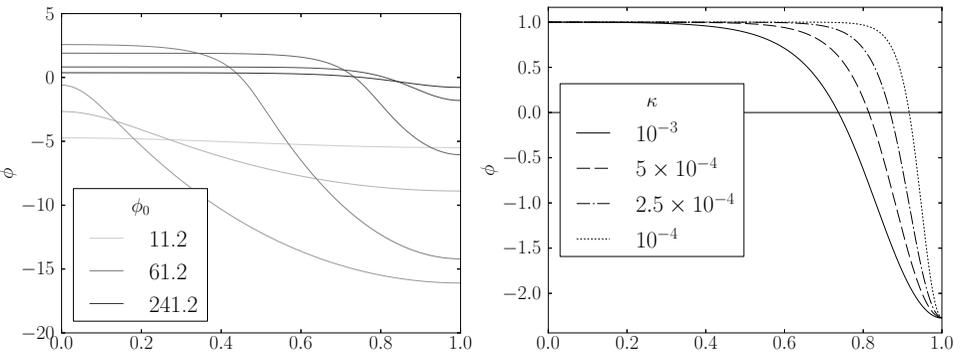
Figure 3. In the case of model (I), $d = 1$, $\delta = 10^{-3}$, we consider various profiles for $x \mapsto \phi(x)$ with $x \in (0, 1) = \Omega$ either (left) as ϕ_0 varies and $\kappa = 5 \times 10^{-4}$, or (right) as κ varies, with $\phi(0) = 1$.

Corollary 8. *Under the assumptions of Lemma 3, for any given $\phi_0 \in \mathbb{R}$, if ϕ is a solution of (10) and (11), then we have that*

$$\phi_-(\phi_0) \leq \phi(x) \leq \phi_+(\phi_0), \quad \forall x \in \Omega.$$

The minimum of $\mathcal{E}_{\phi_0}$ is achieved by a constant function. Moreover, if (12) has only one solution ϕ, then (10) and (11) also have only one solution, which is constant, and $\phi \equiv \phi_- = \phi_+$.

Proof. We simply observe that, according to the definition (13), we have

$$\mathcal{E}_{\phi_0}[\phi] \geq \frac{\delta}{2} \int_{\Omega} |\phi + \phi_0|^2 \, dx - \int_{\Omega} F(\phi) \, dx$$

and critical points of $\phi \mapsto (\delta/2)|\phi + \phi_0|^2 - F(\phi)$ are precisely the constant solutions of (12) with $f = F'$. □

In the cases which are numerically studied in this paper, there is an additional property which is of particular interest.

Proposition 9. *Consider either model (I) or model (II). Then there exists a constant unstable solution only if $\phi_0 \in (\phi_0^-, \phi_0^+)$.*

Proof. This is an easy consequence of the properties of $f = F'$. The details are left to the reader. □

3.6. *A monotonicity result.* For a given $\phi_0 \in \mathbb{R}$, nonmonotone radial functions always have higher energy $\mathcal{E}_{\phi_0}$ than radial monotone functions. We can state this observation as a slightly more general result as follows.

Proposition 10. *Assume that Ω is the unit ball in $\mathbb{R}^d$, $d \geq 2$, and let $G \in W^{1,\infty}(\Omega)$. Then the functional $\mathcal{G}[\phi] := \frac{1}{2} \int_{\Omega} |\nabla \phi|^2 \, dx - \int_{\Omega} G(\phi) \, dx$ is bounded from below and for any radial nonmonotone function $\phi \in C^2(\Omega)$ satisfying (11), with a finite number of critical points, there exists a radial monotone function $\tilde{\phi}$ which satisfies (11) and coincides with ϕ on a neighborhood of 0 such that $\mathcal{G}[\tilde{\phi}] < \mathcal{G}[\phi]$.*

Proof. With a slight abuse of notation, we consider ϕ as a function of $r = |x| \in [0, 1]$ and assume that it is a solution of

$$\phi'' + G'(\phi) = -\frac{d-1}{r} \phi'.$$

Multiplying by ϕ', we find that

$$\frac{d}{dr} \left(\frac{1}{2} \phi'^2 + G(\phi) \right) = -\frac{d-1}{r} \phi'^2 < 0.$$

Unless ϕ is constant, assume that for some $r_0 \in (0, 1)$ we have $\phi'(r_0) = 0$, and let $G_0 := G(\phi(r_0))$. Integrating on (r_0, r), $r > r_0$, we find that

$$\frac{1}{2} \phi'^2 + G(\phi) < G_0 \quad \text{on} \quad (r_0, 1),$$

and then $\frac{1}{2}\phi'^2 - G(\phi) > \phi'^2 - G_0 > -G_0$ on (r_0, R). Hence we have that

$$\frac{\mathcal{G}[\phi]}{|\Omega|} = \int_0^{r_0} \left(\frac{1}{2}\phi'^2 - G(\phi)\right) r^{d-1}\, dr + \int_{r_0}^1 \left(\frac{1}{2}\phi'^2 - G(\phi)\right) r^{d-1}\, dr > \frac{\mathcal{G}[\tilde{\phi}]}{|\Omega|},$$

where $\tilde{\phi}$ is defined by $\tilde{\phi} \equiv \phi$ on $(0, r_0)$ and $\tilde{\phi} \equiv \phi(r_0)$ on $(r_0, 1)$. $\square$

Proposition 10 shows at the ODE level why radial minimizers of the functional $\mathcal{G}$ have to be monotone. It is also preparation for Lemma 16.

Proof of Lemma 6. A constant solution which satisfies (15) cannot be a global minimizer for ϕ_0 fixed. According to Corollary 8, there exists another constant solution under the assumptions of Lemma 6, which incidentally proves that $\phi_-(\phi_0) < \phi_+(\phi_0)$ with the notation of Section 2.2. The fact that there is a constant stable solution with an energy lower than the energy of the unstable solution is a consequence of Proposition 10. $\square$

Summarizing, for a given $\phi_0 \in \mathbb{R}$, only constant solutions are to be considered for the minimization of $\mathcal{E}_{\phi_0}$. However, the relevant problem in terms of modeling is the problem with a mass constraint, at least in view of the evolution problem, and it is not as straightforward as the problem with a fixed Lagrange multiplier.

4. Energy minimizers under mass constraint

4.1. *Existence and qualitative properties of minimizers.* In this section, we assume that $M > 0$ is fixed and consider $\phi_0^M[D] = \phi_0$ uniquely determined by (9). Let us define the functional

$$D \mapsto \mathcal{F}_M[D] := \frac{\kappa}{2} \int_\Omega |\nabla D|^2\, dx + \frac{\delta}{2} \int_\Omega |D|^2\, dx - \int_\Omega F(D - \phi_0^M[D])\, dx.$$

In such a case, ϕ_0 can be seen as a Lagrange multiplier associated with the mass constraint and $\mathcal{F}_M[D] = \mathcal{E}_{\phi_0}[D - \phi_0]$.

Proposition 11. *Assume that F is a continuous function with a subcritical growth. If Ω is bounded with $C^{1,\alpha}$ boundary for some $\alpha > 0$, then for any $M > 0$, the functional $\mathcal{F}_M$ has at least one minimizer $D = \phi + \phi_0$ with $\phi_0 = \phi_0^M[D]$ in $H^1(\Omega)$, which is such that $\mathcal{F}_M[D] = \mathcal{E}_{\phi_0}[\phi]$, and D is of class $C^\infty(\Omega)$ if F is of class C^∞.*

Proof. It is straightforward to check that $\mathcal{F}_M$ has at least one minimizer in $H^1(\Omega)$ because any minimizing sequence converges up to the extraction of subsequences to a minimum $D = \phi + \phi_0$ by compactness and lower semicontinuity. Then ϕ is a critical point of $\mathcal{E}_{\phi_0}$ and regularity is a standard result of elliptic theory (see, for example, [Brezis 2011]) and bootstrapping methods. $\square$

In models (I) and (II), we, respectively, have $|F(\phi)| < 1$ and $F(\phi) \in [0, \log(2) + \max(0, \phi)]$ so the assumptions of Proposition 11 are satisfied. Notice that it is

not implied anymore that minimizers of $\mathcal{F}_M$ under mass constraint are constant functions and hence they might not be minimizers of $\mathcal{E}_{\phi_0}$.

Lemma 12. *The mass of the density associated with nonconstant solutions of* (10) *and* (11) *is bounded away from* 0 *and* $|\Omega|$.

Proof. Any nonconstant solution of (10) and (11) has mass

$$M = \int_\Omega \frac{1}{1 + e^{-\phi}} \, dx$$

associated with its density, according to (8) and (9). Corollary 8 gives the bounds

$$M_-(\phi_0) := \frac{|\Omega|}{1 + e^{-\phi_+(\phi_0)}} \leq M \leq \frac{|\Omega|}{1 + e^{-\phi_-(\phi_0)}} =: M_+(\phi_0).$$

Let $M^{(-)} := \min\{M_-(\phi_0) : \phi_0 \in (\phi_0^-, \phi_0^+)\}$ and $M^{(+)} := \max\{M_+(\phi_0) : \phi_0 \in (\phi_0^-, \phi_0^+)\}$. Since $\phi_0 \mapsto M_\pm(\phi_0)$ is a continuous function on $\mathbb{R}$, we know from Lemma 6 that $(M^{(-)}, M^{(+)})$ is compactly included in $(0, |\Omega|)$. From Lemma 4, we deduce that $M^{(\pm)} = M_\pm(\phi_0^\mp)$. $\qquad\square$

Notice that Lemma 12 proves Theorem 1(i).

Corollary 13. *With the above notation, we have* $0 < M^{(-)} \leq M^{(+)} < 1$ *and minimizers of* $\mathcal{F}_M$ *are constant functions if* $M \in (0, M^{(-)}) \cup (M^{(+)}, 1)$. *There is a subinterval of* $(M^{(-)}, M^{(+)})$ *in which minimizers of* $\mathcal{F}_M$ *are nonconstant functions.*

Whether minimizers of $\mathcal{F}_M$ are constant solutions or not for some $M \in (M^{(-)}, M^{(+)})$ will be investigated numerically. For small masses, or masses close to the maximal mass $|\Omega|$ corresponding to the limit density $\rho = 1$, we can state one more result.

Corollary 14. *Under the assumptions of Lemma 3, with* $M^{(\pm)} \in (0, |\Omega|)$ *defined as above, there is one and only one solution* ϕ *of* (1)–(5), *with mass* $M \in (0, M^{(-)}) \cup (M^{(+)}, |\Omega|)$. *This solution is constant, and given by* $\phi = -\log(|\Omega|/M - 1)$.

4.2. *A partial symmetry result.*

Lemma 15. *Assume that* $d = 2$. *If* Ω *is a disk, minimizers of* $\mathcal{F}_M$ *are symmetric under reflection with respect to a line which contains the origin.*

Proof. The proof of this lemma is inspired by [Lopes 1996]. Assume that Ω is the unit disk centered at the origin and denote by (x_1, x_2) cartesian coordinates in $\mathbb{R}^2$. Let us also define the open upper half-disk $\Omega_+ := \{x \in \Omega : x = (x_1, x_2), x_1 > 0\}$. If ϕ is a minimizer of $\mathcal{F}_M$, we define $\tilde{\phi}$ by $\tilde{\phi}(x_1, x_2) = \phi(|x_1|, x_2)$, so that $\tilde{\phi}$ is symmetric with respect to the line $x_1 = 0$. Up to a rotation, we can assume that Ω_+ accounts for exactly half of the mass, that is, $\int_{\Omega_+} (1 + e^{-\phi})^{-1} \, dx = M/2$, so that

$\int_\Omega (1 + e^{-\tilde\phi})^{-1} \, dx = M$. Then, up to a reflection, we can assume that Ω_+ accounts for at most half of the value of $\mathscr{F}_M$:

$$\frac{\kappa}{2} \int_{\Omega_+} |\nabla\phi|^2 \, dx + \frac{\delta}{2} \int_{\Omega_+} |\phi + \phi_0|^2 \, dx - \int_{\Omega_+} F(\phi) \, dx \le \frac{1}{2} \mathscr{F}_M[\phi].$$

It is then clear that $\tilde\phi$ is a minimizer of $\mathscr{F}_M$ such that the mass constraint (3) is satisfied. As such, $\tilde\phi$ also solves the Euler–Lagrange equations, with the same Lagrange multiplier ϕ_0 because ϕ and $\tilde\phi$ coincide on Ω_+. Then $w := \phi - \tilde\phi$ solves the equation

$$-\kappa \, \Delta w + h w = 0, \quad \text{with} \quad h := \frac{\delta(\phi - \tilde\phi) + F'(\tilde\phi) - F'(\phi)}{\phi - \tilde\phi},$$

on Ω. Since $F \in C^\infty$ and ϕ and $\tilde\phi$ are continuous, h is bounded. According to [Hörmander 1976, Theorem 8.9.1], Hörmander's uniqueness principle applies. Since $w \equiv 0$ on Ω_+, we actually have $w \equiv 0$ on the entire disk Ω, and so $\phi = \tilde\phi$. $\square$

In higher dimensions, when Ω is a ball, the method can be extended and shows the symmetry of the solutions with respect to hyperplanes, thus proving a result of so-called Schwarz foliated symmetry. The method also applies to the functional $\mathscr{E}_{\phi_0}$ with fixed ϕ_0 and shows that a minimizer is radially symmetric, but this is useless as we already know that the minimum is achieved among constant solutions.

4.3. *One-dimensional minimizers are monotone.* A one-dimensional stationary solution solves an autonomous ODE. This has several interesting consequences.

Lemma 16. *Let $d = 1$ and $M > 0$. Then minimizers of $\mathscr{F}_M$ are monotone, either increasing or decreasing.*

Proof. Assume that ϕ is a minimizer of $\mathscr{F}_M$ and $\Omega = (0, 1)$. If ϕ is not monotone, it has a finite number of extremal points $0 = r_0 < r_1 < \cdots < r_N = 1$ for some $N > 1$. By uniqueness of the solution of the initial value problem, with $\phi(r_i)$ given and $\phi'(r_i) = 0$, we conclude that $\phi(r_i - s) = \phi(r_i + s)$ as long as $0 \le r_i - s$ and $r_i + s \le 1$, so that $r_i = i/N$, that is, ϕ is $1/N$-periodic. With $\tilde\phi(r) := \phi(r/N)$, $r \in (0, 1)$, we find that

$$\int_0^1 |\tilde\phi'|^2 \, dr = \frac{1}{N} \int_0^{1/N} |\phi'|^2 \, dr = \frac{1}{N^2} \int_0^1 |\phi'|^2 \, dr < \int_0^1 |\phi'|^2 \, dr,$$

thus proving that $\mathscr{E}_{\phi_0}[\tilde\phi] < \mathscr{E}_{\phi_0}[\phi]$ while $\int_0^1 (1 + e^{-\tilde\phi}) \, dr = \int_0^1 (1 + e^{-\phi}) \, dr$, a contradiction. $\square$

From the scaling in the above proof, it is now clear that *all* nonmonotone one-dimensional solutions can be built from monotone ones by symmetrizing them with respect to their critical points, duplicating them, and scaling them. The intuitive

idea is simple but giving detailed statements is unnecessarily complicated, so we will focus on monotone, or *one-plateau*, solutions.

5. A Lyapunov functional

In the case of model (II), let us consider the functional

$$\mathscr{L}[\rho, D] := \int_\Omega [\rho \log \rho + (1-\rho) \log(1-\rho) - \rho D] \, dx + \frac{\kappa}{2} \int_\Omega |\nabla D|^2 \, dx + \frac{\delta}{2} \int_\Omega D^2 \, dx.$$

Proposition 17. *The functional $\mathscr{L}$ is a* Lyapunov functional *for model (II) and if (ρ, D) is a solution of (1)–(5) and (7), then*

$$\frac{d}{dt} \mathscr{L}[\rho(t, \cdot), D(t, \cdot)] = -\int_\Omega \frac{|\nabla \rho - \rho(1-\rho)\nabla D|^2}{\rho(1-\rho)} \, dx - \int_\Omega |-\kappa \Delta D + \delta D - \rho|^2 \, dx \le 0.$$

As a consequence, any critical point of $\mathscr{L}$ under the mass constraint (3) is a stationary solution of (1)–(5) and (7), and any solution converges to a stationary solution. If Ω is a ball and if the initial datum is radial, then the limit is a radial stationary solution.

Proof. An elementary computation shows that

$$\frac{d}{dt} \mathscr{L}[\rho(t, \cdot), D(t, \cdot)] = -\int_\Omega \left[\log\left(\frac{\rho}{1-\rho}\right) - D\right] D_t \, dx - \int_\Omega (-\kappa \Delta D + \delta D - \rho) D_t \, dx$$

and the expression of

$$\frac{d}{dt} \mathscr{L}[\rho(t, \cdot), D(t, \cdot)]$$

follows from (1)–(5). Let $\rho_n(t, x) := \rho(t+n, x)$ and $D_n(t, x) := D(t+n, x)$. Since $\mathscr{L}$ is bounded from below, we have that

$$\lim_{n\to\infty} \int_0^1 \left(\int_\Omega \frac{|\nabla \rho_n - \rho_n(1-\rho_n)\nabla D_n|^2}{\rho_n(1-\rho_n)} \, dx + \int_\Omega |-\kappa \Delta D_n + \delta D_n - \rho_n|^2 \, dx\right) dt = 0,$$

which proves that (ρ_n, D_n) strongly converges to a stationary solution. Other details of the proof are left to the reader. $\qquad\square$

Proposition 18. *Let $M > 0$ and consider model (II). For any $D \in H^1(\Omega)$, let ϕ_0 be the unique real number determined by the mass constraint (9). Then for any nonnegative $\rho \in L^1(\Omega)$ satisfying the mass constraint (3), we have*

$$\mathscr{L}[\rho, D] \ge \mathscr{E}_{\phi_0}[D - \phi_0],$$

where equality holds if and only if ρ is given by (8), that is, $\rho = 1/(1 + e^{-\phi})$, with $\phi = D - \phi_0$. As a consequence, for any minimizer (ρ, D) of $\mathscr{L}$ satisfying (3), ρ is given by (8) with $\phi = D - \phi_0$, ϕ_0 satisfying the mass constraint (9), and ϕ is a minimizer of $\mathscr{E}_{\phi_0}[\phi] = \mathscr{L}[\rho, D]$.

Proof. We only need to notice that the minimum of $\mathcal{L}[\rho, D]$ with respect to ρ under the mass constraint (3) satisfies

$$\log\left(\frac{\rho}{1-\rho}\right) = D - \phi_0.$$

The completion of the proof follows from elementary computations which are left to the reader.   □

6. The linearized evolution operator

6.1. *Dynamical instability of constant solutions.* Assume that (ρ, D) is a stationary solution of (1)–(5). Because of (8) and (9), the solution is fully determined by D. Let us consider a time-dependent perturbed solution of the form $(\rho + \varepsilon u, D + \varepsilon v)$. Up to higher-order terms, u and v are solutions of the linearized system

$$\begin{cases} u_t = \nabla \cdot (\nabla u - (1 - 2\rho)u\nabla D - \rho(1 - \rho)\nabla v), \\ v_t = \kappa \Delta v - \delta v + h(\rho)u, \end{cases} \tag{16}$$

with $h(\rho) = 1 - 2\rho$ in the case of model (I) and $h(\rho) = 1$ in the case of model (II). For later use, we introduce the notation H_D for the linear operator corresponding to the right-hand side, so we write

$$(u, v)_t = H_D(u, v) = \begin{pmatrix} H_D^{(1)}(u, v) \\ H_D^{(2)}(u, v) \end{pmatrix},$$

with $\quad \begin{cases} H_D^{(1)}(u, v) = \nabla \cdot \left[\rho(1 - \rho)\nabla\left(\dfrac{u}{\rho(1-\rho)} - v\right)\right], \\ H_D^{(2)}(u, v) = \kappa \Delta v - \delta v + h(\rho)u. \end{cases}$

Dynamical instability of constant solutions can be studied along the lines of [Burger et al. 2011]. Let us state a slightly more general result. We are interested in finding the lowest possible μ in the eigenvalue problem

$$-H_D(u, v) = \mu(u, v), \tag{17}$$

where $H_D^{(1)}(u, v)$ now takes a simplified form, using the fact that ρ is a constant:

$$H_D^{(1)}(u, v) = \Delta u - \rho(1 - \rho)\Delta v.$$

The condition $\mu < 0$ provides a dynamically unstable mode. As in Section 3.3, let us denote by $(\lambda_n)_{n\in\mathbb{N}}$ the sequence of all eigenvalues of $-\Delta$ with homogeneous Neumann boundary conditions, counted with multiplicity, and by $(\phi_n)_{n\in\mathbb{N}}$ an associated sequence of eigenfunctions. If $u = \sum_{n\in\mathbb{N}} \alpha_n \phi_n$ and $v = \sum_{n\in\mathbb{N}} \beta_n \phi_n$,

problem (17) can be decomposed into

$$-\lambda_n \alpha_n + \rho(1-\rho)\lambda_n \beta_n = -\mu_n \alpha_n,$$
$$-\kappa \lambda_n \beta_n - \delta \beta_n + h(\rho)\alpha_n = -\mu_n \beta_n,$$

for any $n \in \mathbb{N}$, that is,

$$(\mu_n - \lambda_n)\alpha_n + \rho(1-\rho)\lambda_n \beta_n = 0,$$
$$h(\rho)\alpha_n + (\mu_n - \kappa\lambda_n - \delta)\beta_n = 0,$$

which has nontrivial solutions α_n and β_n if and only if the discriminant condition

$$(\mu_n - \lambda_n)(\mu_n - \kappa\lambda_n - \delta) - \rho(1-\rho)h(\rho)\lambda_n = 0$$

is satisfied. This determines μ_n for any $n \in \mathbb{N}$, and the spectrum of H_D is then given by $(\mu_n)_{n \in \mathbb{N}}$. Collecting these observations, we can state the following result.

Proposition 19. *With the above notation,* $\inf_{n \geq 1} \mu_n < 0$ *if and only if (15) holds.*

Proof. The discriminant condition can be written as

$$\mu_n^2 - [(\kappa+1)\lambda_n + \delta]\mu_n + \lambda_n[\kappa\lambda_n + \delta - \rho(1-\rho)h(\rho)] = 0,$$

so that there is a negative root if $\lambda_n(\kappa\lambda_n + \delta - \rho(1-\rho)h(\rho)) < 0$. Since $(\lambda_n)_{n \in \mathbb{N}}$ is nondecreasing and $\lambda_0 = 0$, there is at least one negative eigenvalue for (17) if the above condition is satisfied with $n = 1$. $\qquad\square$

In other words, the dynamical instability of the constant solutions implies their variational instability. As we shall see numerically, variational and dynamical instability are not equivalent for plateau-like solutions.

Notice that $\lambda_0 = 0$ must still be excluded, as it corresponds to the direction generated by constants. Because of (15) we can ensure that the perturbation has zero average. This will be discussed further below, in the general case of a stationary solution.

6.2. *Variational criterion.* In the case of model (II), we can look at the Lyapunov functional $\mathcal{L}$ and linearize it around a stationary solution (ρ, D). Let

$$L_D[u, v] := \lim_{\varepsilon \to 0} \frac{\mathcal{L}[\rho + \varepsilon u, D + \varepsilon v] - \mathcal{L}[\rho, D]}{2\varepsilon^2}.$$

A simple computation shows that

$$L_D[u, v] = \int_\Omega \left(\frac{u^2}{2\rho(1-\rho)} - uv \right) dx + \frac{\kappa}{2} \int_\Omega |\nabla v|^2 \, dx + \frac{\delta}{2} \int_\Omega v^2 \, dx.$$

With E_ϕ defined by (14), let

$$\Lambda := \inf_{\substack{\int_\Omega v\rho(1-\rho)\,dx=0 \\ v\not\equiv 0}} \frac{\int_\Omega v(E_\phi v)\,dx}{\int_\Omega v^2\,dx}, \tag{18}$$

with $\phi = D - \phi_0$, and ϕ_0 satisfying (9).

Lemma 20. *Let $M > 0$ and consider model (II) only. If (ρ, D) satisfies (3) and is such that ρ is given by (8) with $\phi = D - \phi_0$ and ϕ_0 determined by (9), then*

$$\Lambda = 2 \inf_{\substack{\int_\Omega v\rho(1-\rho)\,dx=0 \\ \int_\Omega v^2\,dx=1}} L_D[u, v].$$

As a consequence, if (ρ, D) is a local minimizer of $\mathcal{L}$ under the mass constraint (3), then Λ is nonnegative.

Proof. We notice that $L_D[u, v] = \int_\Omega v(E_\phi v)\,dx$ holds true as soon as $u = v\rho(1-\rho)$, with ρ given by (8). In particular, this is the case if (ρ, D) is a local minimizer of $\mathcal{L}$. With v fixed, an optimization of $L_D[u, v]$ with respect to u shows that $u = v\rho(1-\rho)$. When (ρ, D) is a local minimizer of $\mathcal{L}$, it is straightforward to check that $L_D[u, v]$ cannot be negative.   □

6.3. *Entropy-entropy production.* Along the linearized flow (16), we have

$$\frac{d}{dt} L_D[u(t, \cdot), v(t, \cdot)] = -2I_D[u(t, \cdot), v(t, \cdot)], \tag{19}$$

where

$$I_D[u, v] := \frac{1}{2} \int_\Omega \rho(1-\rho) \left| \nabla \left(\frac{u}{\rho(1-\rho)} - v \right) \right|^2 dx + \frac{1}{2} \int_\Omega |-\kappa \Delta v + \delta v - u|^2\,dx.$$

Let us define the bilinear form

$$\langle (u_1, v_1), (u_2, v_2) \rangle_D$$
$$= \int_\Omega \left(\frac{u_1 u_2}{\rho(1-\rho)} - (u_1 v_2 + u_2 v_1) \right) dx + \kappa \int_\Omega \nabla v_1 \cdot \nabla v_2\,dx + \delta \int_\Omega v_1 v_2\,dx,$$

which is such that

$$2L_D[u, v] = \langle (u, v), (u, v) \rangle_D.$$

Lemma 21. *Consider model (II) only and assume that (ρ, D) is a local minimizer of $\mathcal{L}$ under the mass constraint (3). On the orthogonal of the kernel of E_ϕ with $\phi = D - \phi_0$, ϕ_0 satisfying (9), $\langle \cdot, \cdot \rangle_D$ is a scalar product and H_D is a self-adjoint operator with respect to $\langle \cdot, \cdot \rangle_D$. Moreover, if (u, v) is a solution of (16), then*

$$\frac{d}{dt} L_D[u, v] = -2I_D[u, v] = \langle (u, v), H_D(u, v) \rangle_D \le 0.$$

As a consequence, on the orthogonal of the kernel of E_ϕ, $(0, 0)$ *is the unique stationary solution of* (16) *and any solution with initial datum in the orthogonal of the kernel converges to* $(0, 0)$.

With a slight abuse of notation, we have denoted by the *kernel of* E_ϕ the set $\{(u, v) : v \in \text{Ker}(E_\phi)\}$.

Proof. The positivity of I_D is a consequence of the definition and self-adjointness results from the computation

$$-\langle(u_1, v_1), H_D(u_2, v_2)\rangle_D = \int_\Omega \rho(1-\rho)\nabla\left(\frac{u_1}{\rho(1-\rho)} - v_1\right)\cdot\nabla\left(\frac{u_2}{\rho(1-\rho)} - v_2\right) dx$$

$$+ \int_\Omega (-\kappa\Delta v_1 + \delta v_1 - u_1)(-\kappa\Delta v_2 + \delta v_2 - u_2) dx. \quad \square$$

Note that one has to take special care of the kernel of H_D. If (ρ_M, D_M) is a stationary solution of (1)–(4) depending differentiably on the mass parameter M, it is always possible to differentiate ρ_M and D_M with respect to M and get a nontrivial element in the kernel of H_D. However, it is not guaranteed that this element generates the kernel of E_ϕ, and, although not observed numerically, it cannot be excluded that secondary bifurcations occur on branches of plateau-like solutions.

If (ρ, D) is a stationary solution of (1)–(5), we can of course still consider $I_D[u, v]$, and its sign determines whether (ρ, D) is dynamically stable or not. In this paper we are interested in the evolution according to the nonlinear flow given by (1)–(5). The fundamental property of mass conservation (3) can still be observed at the level of the linearized equations (16). The reader is invited to check that any classical solution of (16) is indeed such that

$$\frac{d}{dt}\int_\Omega u(t, x)\, dx = 0,$$

and it makes sense to impose $\int_\Omega u\, dx = 0$ at $t = 0$. If we linearize the problem at a stationary solution given by (8), it also makes sense to consider the constraint $\int_\Omega v\rho(1-\rho)\, dx = 0$.

6.4. *Dynamic criterion.* After these preliminary observations, we can define two notions of stability. We shall say that a critical point ϕ of $\mathcal{E}_{\phi_0}$ is *variationally stable* (resp. unstable) if and only if $\Lambda > 0$ (resp. $\Lambda < 0$), where Λ is defined by (18). Alternatively, we shall say that a stationary solution (ρ, D) of (1)–(5) is *dynamically stable* (resp. unstable) if and only if

$$\inf_{\substack{\int_\Omega u\, dx = 0 \\ \int_\Omega v^2\, dx = 1}} L_D[u, v]$$

is positive (resp. negative) in the case of model (II). The operator H_D being self-adjoint, dynamical stability means *variational stability* of $\mathscr{L}$ on the product space, once mass constraints are taken into account. Most of the remainder of this section is devoted to this issue.

For model (I) we can extend the notion of dynamical stability (resp. instability) by requiring that $\inf\{\mathrm{Re}(\lambda) : \lambda \in \mathrm{Spectrum}(H_D)\}$ is positive (resp. negative). However, in the case of model (I), notions of dynamical and variational instability are not so well related, as we shall see in Section 7.

Let us start with the following observation. In the cases of models (I) and (II), the kernel of the operator E_ϕ associated with the linearized energy functional and defined by (14) determines a subspace of the kernel of H_D.

Lemma 22. *Let $\phi_0 \in \mathbb{R}$ and assume that ϕ is a critical point of $\mathscr{E}_{\phi_0}$. Then ρ given by (8) and $D = \phi + \phi_0$ provides a stationary solution of (1)–(5). If v is in the kernel of E_ϕ, then (u, v) is in the kernel of H_D if $u = \rho(1 - \rho)v$.*

Proof. Using (14), it is straightforward to check that $0 = E_\phi v = H_D^{(2)}(u, v)$ if $v \in \mathrm{Ker}(E_\phi)$. Then

$$H_D^{(1)}(u, v) = \nabla \cdot \left[\rho(1 - \rho)\nabla \left(\frac{u}{\rho(1 - \rho)} - v \right) \right] = 0$$

because of the special choice $u = \rho(1 - \rho)v$. $\qquad\square$

Since (1) preserves the mass, it makes sense to impose $\int_\Omega u \, dx = 0$. This also suggests considering the constraint $\int_\Omega \rho(1 - \rho)v \, dx = 0$, which has already been taken into account in (18). Let us give some more precise statements, in the case of model (II). First we can state a more precise version of Lemma 20. Let us define

$$\Lambda_1 := 2 \inf_{\substack{\int_\Omega u \, dx = 0 \\ \int_\Omega v^2 \, dx = 1}} L_D[u, v].$$

Lemma 23. *Let $M > 0$. Consider model (II) only and assume that (ρ, D) is a critical point of $\mathscr{L}$ under the mass constraint (3). With $\phi = D - \phi_0$ where ϕ_0 is the unique real number determined by (9), consider Λ defined by (18). Then we have $\Lambda_1 \leq \Lambda$. If either $\Lambda < \delta$ or $\Lambda_1 < \delta$, then we have $\Lambda = \Lambda_1$.*

Proof. If (ρ, D) is a critical point of $\mathscr{L}$, the analysis of Section 2.1 shows that ρ is given by (8) with $\phi = D - \phi_0$ and ϕ_0 determined by (9). Consider first the minimization problem

$$\inf_{\substack{\int_\Omega v\rho(1-\rho) \, dx = 0 \\ \int_\Omega v^2 \, dx = 1}} L_D[u, v].$$

As in Lemma 20, optimization with respect to u shows that $u = v\rho(1 - \rho)$, and it is then straightforward to get that $2L_D[u, v] = \int_\Omega v(E_\phi v) \, dx = \Lambda$. Additionally,

we know that v solves the Euler–Lagrange equation

$$E_\phi v = -\kappa \Delta v + \delta v - v\rho(1-\rho) = \Lambda v - \mu\rho(1-\rho), \qquad (20)$$

for some Lagrange multiplier μ, and we have $\int_\Omega u \, dx = \int_\Omega v\rho(1-\rho) \, dx = 0$. This proves that $\Lambda_1 \le \Lambda$.

Now, consider a minimizer (u, v) for Λ_1. We find that

$$u = (v - \bar{v})\rho(1-\rho), \quad \text{with} \quad \bar{v} := \frac{\int_\Omega v\rho(1-\rho) \, dx}{\int_\Omega \rho(1-\rho) \, dx}.$$

Moreover, v solves the Euler–Lagrange equation

$$-\kappa \Delta v + \delta v - v\rho(1-\rho) = \Lambda_1 v - \bar{v}\rho(1-\rho). \qquad (21)$$

Hence we have found that

$$2\mathrm{L}_D[u, v] = \kappa \int_\Omega |\nabla v|^2 \, dx + \delta \int_\Omega v^2 \, dx - \int_\Omega \rho(1-\rho)|v - \bar{v}|^2 \, dx,$$

so that

$$
\begin{aligned}
\Lambda_1 - \delta &= \inf_{v \ne 0} \frac{\kappa \int_\Omega |\nabla v|^2 \, dx - \int_\Omega \rho(1-\rho)|v - \bar{v}|^2 \, dx}{\int_\Omega v^2 \, dx} \\
&= \inf_{\substack{\int_\Omega v\rho(1-\rho) \, dx=0 \\ v \ne 0, c \in \mathbb{R}}} \frac{\kappa \int_\Omega |\nabla v|^2 \, dx - \int_\Omega \rho(1-\rho)v^2 \, dx}{\int_\Omega v^2 \, dx + c^2} \\
&= \inf_{\substack{\int_\Omega v\rho(1-\rho) \, dx=0 \\ v \ne 0}} \frac{\kappa \int_\Omega |\nabla v|^2 \, dx - \int_\Omega \rho(1-\rho)v^2 \, dx}{\int_\Omega v^2 \, dx},
\end{aligned}
$$

where the last equality holds under the condition that either $\Lambda < \delta$ or $\Lambda_1 < \delta$. Hence we have shown that $\Lambda_1 - \delta = \Lambda - \delta$, which concludes the proof. $\qquad \square$

Remark 24. With no constraint, it is straightforward to check that δ is an eigenvalue of H_D, and $(u, v) = (0, 1)$ an eigenfunction. Hence, as soon as $\Lambda_1 < \delta$, we have that $\int_\Omega u \, dx = 0$ if (u, v) is a minimizer for Λ_1, because of (21). This justifies why the condition of either $\Lambda < \delta$ or $\Lambda_1 < \delta$ enters into the statement of Lemma 23.

In the case of model (II), we can get a bound on the growth of the unstable mode.

Corollary 25. *Consider model* (II) *only and assume that* (ρ, D) *is a critical point of* $\mathscr{L}$ *under the mass constraint* (3)*. If* Λ *defined by* (18) *is negative, then we have*

$$\inf_{\substack{\int_\Omega u \, dx=0 \\ \int_\Omega v^2 \, dx=1}} \frac{\mathrm{I}_D[u, v]}{\mathrm{L}_D[u, v]} \le \Lambda$$

and the growth rate of the most unstable mode of (16) *is at least* $2|\Lambda|$.

Proof. Consider a function v given by (20) with $\int_\Omega v\rho(1-\rho)\,dx = 0$, $\int_\Omega v^2\,dx = 1$, $u = v\rho(1-\rho)$, and (u,v) taken as a test function. Then

$$
\begin{aligned}
\frac{I_D[u,v]}{L_D[u,v]} &= \frac{\int_\Omega (E_\phi v)^2\,dx}{\int_\Omega v(E_\phi v)\,dx} \\
&= \frac{\int_\Omega (\Lambda v - \mu\rho(1-\rho))^2\,dx}{\int_\Omega (\Lambda v - \mu\rho(1-\rho))v\,dx} = \Lambda + \frac{\mu^2}{\Lambda}\int_\Omega \rho^2(1-\rho)^2\,dx \leq \Lambda.
\end{aligned}
$$

Using (19), if $L_D[u,v]$ is negative, then we get

$$
\frac{d}{dt}L_D[u,v] \leq -2\Lambda L_D[u,v],
$$

thus proving that $L_D[u,v](t) \leq L_D[u,v](0)e^{2|\Lambda|t}$ for any $t \geq 0$. $\qquad\square$

The result of Corollary 25 on the most unstable mode can be rephrased in terms of standard norms. By definition of L_D, we get that

$$
\int_\Omega (u^2 + v^2)\,dx \geq 2\int_\Omega uv\,dx \geq 2|L_D[u,v]| \geq 2|L_D[u,v](0)|e^{2|\Lambda|t},
$$

for any $t \geq 0$.

Summarizing, we have shown the following result.

Theorem 26. *Let $M > 0$ and consider the case of model* (II)*. Assume that (ρ, D) is a stationary solution of* (1)–(5) *such that* (3) *is satisfied and let $\phi = D - \phi_0$, with ϕ_0 satisfying* (9)*. Then the following properties hold true.*

(i) *Neither dynamical instability nor variational instability can occur if (ρ, D) is a local minimizer of $\mathcal{L}$ under the mass constraint* (3) *or, equivalently, if ϕ is a local minimizer of $\mathcal{E}_{\phi_0}$ such that* (3) *and* (8) *hold.*

(ii) *If (ρ, D) is a local minimizer of $\mathcal{L}$ under the mass constraint* (3)*, then any solution (u, v) of* (16) *converges towards* $(0, 0)$ *when the initial datum is assumed to be in the orthogonal of the kernel of H_D and with sufficiently low energy.*

(iii) *Dynamical stability implies variational stability.*

(iv) *Variational instability and dynamical instability are equivalent and, with the above notation, $\Lambda_1 = \Lambda$.*

On the contrary, no clear relation between variational and dynamical (in)stability is known in the case of model (I), except the result of Lemma 22, which is not so easy to use from a numerical point of view.

7. Numerical results

Let us summarize our findings on radial stationary solutions of (1)–(5), with parameters δ and κ in the range discussed in Section 3.4, when Ω is the unit ball in $\mathbb{R}^d$, with $d = 1$ or $d = 2$. Our results deal with either model (I) or model (II), defined respectively by (6) and (7), as follows:

(i) We compute the branches of monotone, nonconstant, radial solutions that bifurcate from constant solutions for the two models, in dimensions $d = 1$ and $d = 2$.

(ii) We study the variational and dynamical stability of these solutions. The two notions coincide for model (II), which is partially explained with the help of the Lyapunov functional.

(iii) Dynamical stability holds up to the turning point of the branch when it is parametrized by the mass for model (II) in dimensions $d = 1$ and $d = 2$. This is also true in dimension $d = 1$ for model (I).

(iv) In dimension $d = 1$, the variational stability of the branch of monotone, nonconstant solutions is more restrictive than the dynamical stability in the case of model (I).

Before entering into the details, let us observe that bifurcation diagrams are more complicated in dimension $d = 2$ than for $d = 1$, and that the lack of a Lyapunov functional makes the study of model (I) significantly more difficult.

All computations are based on the shooting method presented in Proposition 2. This allows us to find all radially symmetric stationary solutions, as the range of parameter a for which solutions exist is bounded according to Corollary 8. Hence we are left with a single ordinary differential equation, which can be solved using standard numerical methods. Because of the smallness of the parameter κ, the shooting criterion $\varphi'_a(1) = 0$ has a rather stiff dependence on a. This makes directly finding all zeros of the criterion for a given ϕ_0 difficult, so in practice we use perturbation and continuation methods to parametrize the whole branch of monotone, plateau-like solutions.

The computation of the spectrum of the linearized evolution operator (16) is done using a basis of cosines, normalized and scaled to meet the boundary conditions. This allows for fast decomposition of the coefficients by FFT. In the case $d = 2$, such a basis in not orthogonal, which is taken into account using a mass matrix during diagonalization. In cases where the constraints cannot be enforced directly at the basis level, a Rayleigh quotient minimization step is performed, on the orthogonal of the constrained space.

Numerical computations have been made entirely using the freely available NumPy and SciPy Python libraries. These make use of reference numerical libraries LAPACK and odepack.

We start by considering constant solutions. We make use of the notation of Section 2.2.

Let us comment on the plots of Figure 2.

$\boxed{1}$ The first turning point: $\phi_0 = \phi_0^-$, on the branch of constant solutions. For lower values of ϕ_0, there is only one constant solution $\phi \equiv \phi(0)$, which converges to $+\infty$ as $\phi_0 \to -\infty$.

$\boxed{2}$ and $\boxed{3}$ Nonconstant solutions bifurcate from constant solutions, which are unstable in the corresponding interval for ϕ_0. The solutions of the two branches correspond to monotone solutions, either increasing or decreasing, and always bounded from above and below by constant solutions.

$\boxed{4}$ Second turning point: $\phi_0 = \phi_0^+$, on the branch of constant solutions. For higher values of ϕ_0, there is only one constant solution $\phi \equiv \phi(0)$, which converges to $-\infty$ as $\phi_0 \to +\infty$.

The dependence of plateau-like solutions on parameters ϕ_0 and κ is shown in Figure 3.

Next we consider monotone, plateau-like solutions. In Figures 4 and 5, the shaded region corresponds to masses for which constant solutions are unstable.

Dynamical and variational stability criteria and their interplay are a tricky issue, especially in the case of model (I), in which we have no theoretical framework to relate the two notions. See Figure 6.
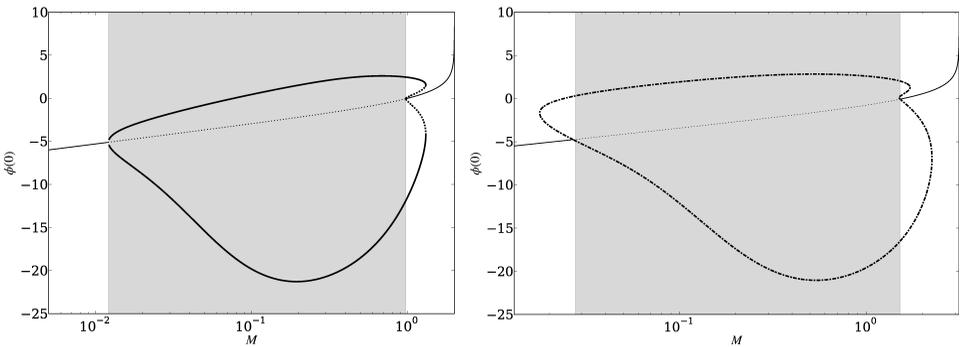

Figure 4. Model (I), with $\kappa = 5 \times 10^{-4}$ and $\delta = 10^{-3}$. Thin lines represent constant solutions and bold lines plateau-like solutions. For readability purposes we use a logarithmic scale for the mass. The dotted part of each branch shows where solutions are dynamically unstable. For $d = 1$ (left) and $d = 2$ (right).
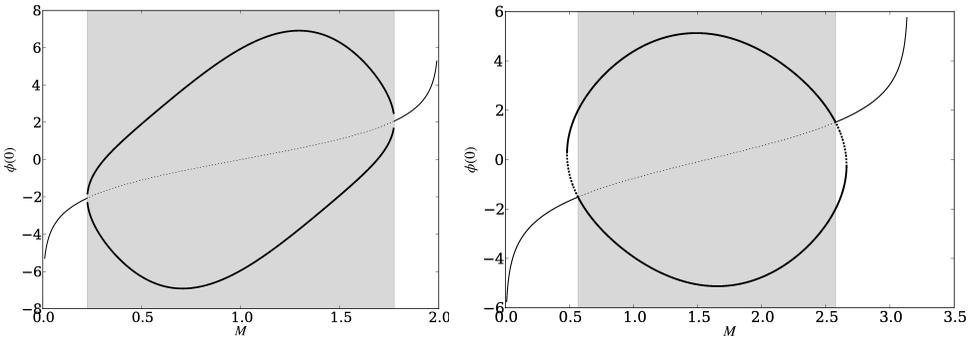
Figure 5. Model (II), with $\kappa = 10^{-2}$ and $\delta = 10^{-3}$. Thin lines represent constant solutions and bold lines plateau-like solutions. The dotted part of each branch shows where solutions are dynamically unstable. For $d = 1$ (left) and $d = 2$ (right).

Stationary solutions are critical points of $\mathscr{E}_{\phi_0}$. It is therefore interesting to determine whether they are minima or not, either for fixed values of ϕ_0 or for fixed values of M, which makes more sense from the dynamical point of view. However, it is only in the case of model (II) that minimizers of $\mathscr{E}_{\phi_0}$ are also minimizers of $\mathscr{L}$, and therefore dynamically and variationally stable. See Figures 7 and 8.

Finally in the case of model (II), we can check that dynamical and variational stability are compatible; see Figure 9.

8. Concluding remarks

Model (II) is the (formal) gradient flow of the Lyapunov functional $\mathscr{L}$ with respect to a distance corresponding to the Wasserstein distance for ρ and an L^2 distance for D (see [Blanchet and Laurençot 2013; Blanchet et al. 2015; Laurençot and Matioc 2013] for further considerations in this direction). Critical points of $\mathscr{L}$ are stationary solutions for the system. They attract all solutions of the evolution equation and the infimum of $\mathscr{L}$ is achieved by a monotone function, which is therefore either a plateau solution or a constant solution. When $d = 1$, numerics, at least for the values of the parameters we have considered, show that plateau solutions exist only in the range in which constant solutions are unstable and are uniquely defined in terms of the mass. However, when $d = 2$, the range for dynamically stable plateau solutions is larger than the range (in terms of the mass) of constant unstable solutions under radial perturbation. Infima of $\mathscr{L}$ and $\mathscr{E}_{\phi_0}$ actually coincide. Consistent with our analysis, we find that the linearized evolution operator around minimizing solutions has only positive eigenvalues. Moreover, this operator is self-adjoint in the norm corresponding to the quadratic form given by the second variation of $\mathscr{L}$ around a
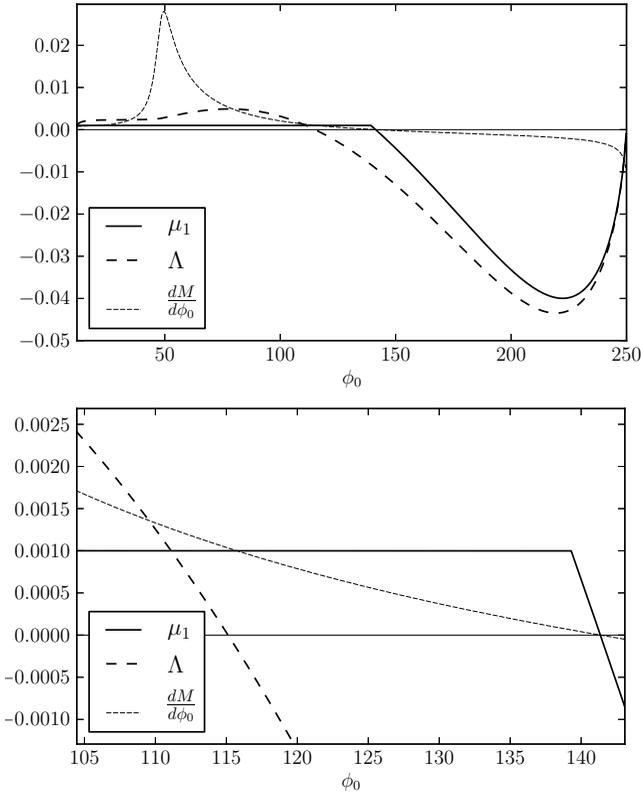
Figure 6. Model (I) with $d = 1$. We numerically compare the criteria for variational and dynamical instability along the branch of monotone, nonconstant solutions. When $dM/d\phi_0$ changes sign, this means that the branch has a turning point when plotted in terms of M. We observe that this turning point corresponds to the loss of dynamical stability, while variational stability is lost for smaller values of ϕ_0 along the branch; see in particular the enlargement (bottom). Here μ_1 corresponds to the lowest value of $\mathrm{Re}(\langle (u, v), -\mathrm{H}_D(u, v) \rangle)$ under the constraints $\langle (u, v), (u, v) \rangle = 1$ and $\int_\Omega u \, dx = 0$, where $\langle \cdot, \cdot \rangle_D$ denotes the standard scalar product.

minimizer. Hence, when $d = 2$, we observe the existence of multiple stable (under radial perturbation) stationary solutions.

In the case of *model (I)*, no Lyapunov functional is available, to our knowledge. Still, all stationary states are characterized as critical points of $\mathcal{E}_{\phi_0}$ and obtained (as long as they are radially symmetric) using our shooting method. In dimension $d = 1$, the structure of the set of solutions is not as simple as in model (II), and this can be explained by the frustration due to the $\rho(1 - \rho)$ term in the equation
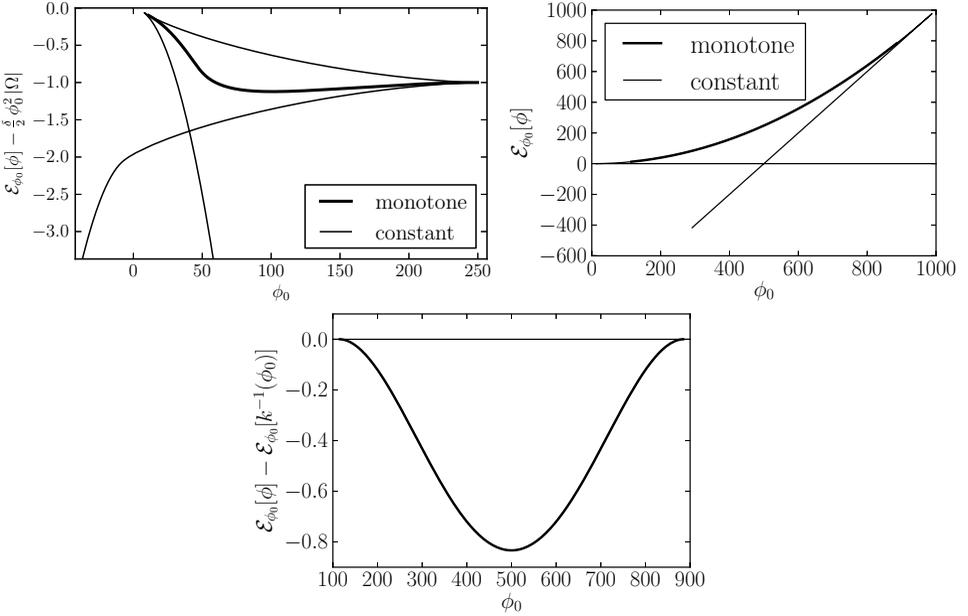
Figure 7. Energy is represented as a function of ϕ_0 for constant and monotone (either increasing or decreasing) solutions. Here we assume $d = 1$. For model (I) (upper left), the energy $\mathcal{E}_{\phi_0}$ is shifted by $(\delta/2)\phi_0^2|\Omega|$. For model (II) (upper right), nonconstant solutions (the upper curve) are indistinguishable from a branch of constant solutions. Details for model (II) (bottom) show the difference of the energies of the constant and nonconstant solutions (under appropriate restrictions on ϕ_0).
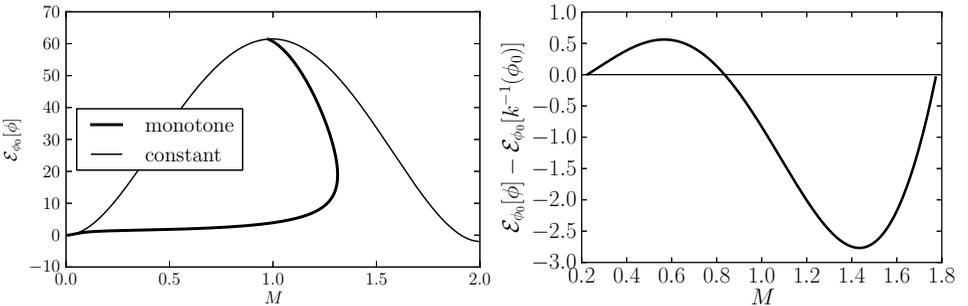
Figure 8. For any given mass, there is exactly one constant solution. Hence minimizers of $\mathcal{F}_M[D] = \mathcal{E}_{\phi_0}[\phi + \phi_0]$ with $\phi_0 = \phi_0^M[D]$ for masses M in a certain range are not constant. Model (I) with $d = 1$ (left) and model (II) with $d = 1$ (right). These minimizers are also minimizers of the Lyapunov functional and therefore dynamically stable (ϕ_0 is restricted to an appropriate range).
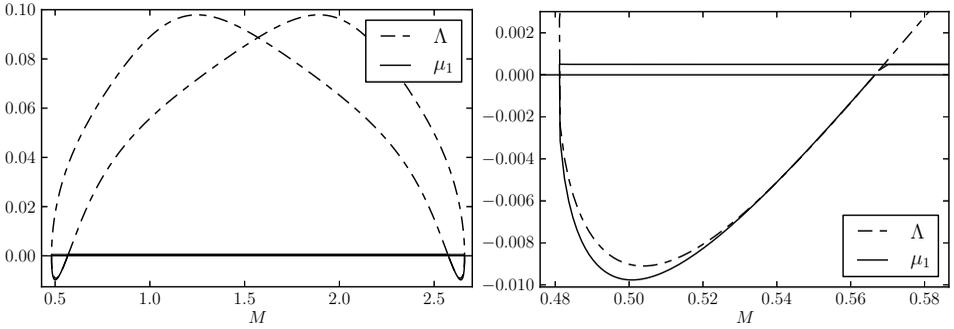
Figure 9. Model (II) with $d = 2$. The solution of $M \mapsto \Lambda$ (left), where, for each M, we compute the two monotone plateau-like solutions, and then Λ according to (18). Hence $\Lambda < 0$ means that the solution is variationally unstable under the mass constraint. Detail is shown (right). Here μ_1 corresponds to the lowest value of $\langle (u, v), -H_D(u, v) \rangle$ under the constraints $\langle (u, v), (u, v) \rangle = 1$ and $\int_\Omega u \, dx = 0$.

for D. Numerically, when $d = 1$, we observe that monotone plateau solutions are uniquely defined and dynamically stable in the range where constant solutions are dynamically unstable. However, when $d = 1$, we also have a range in which both types of solutions are dynamically stable, which means that the system has no global attractor. We do not even know whether stationary solutions attract all solutions of the evolution problem or not.

To give a simple picture of the physics involved in the two models of crowd modeling studied in this paper, we may use the following image. The potential D defines the *strategy* of the individuals. It takes into account the source term (the density ρ in the case of model (II) and $\rho(1 - \rho)$ in the case of model (I)) to determine a preferred direction. Because it is governed by a parabolic equation, it takes the value of the source term into account not only at instant t, but also in the past, which means that there is a memory effect. Of course, the recent past receives a larger weight, and actually two mechanisms are at work updating the system: a local damping, with time scale determined by δ, and a diffusion term (the position of the source term gets lost over a long time range), with a time scale governed by κ. Both coefficients being small, the time scale (that is, the memory of the system) is long compared to the time scale for ρ.

As far as ρ is concerned, the diffusion accounts for random effects while the drift is tempered by some *tactical* term, which tries to avoid densely populated areas, and is taken into account by means of the $(1 - \rho)$ term in the drift.

In the case of model (II) the strategy defined by the source term is simple: individuals want to aggregate in high ρ densities. In the case of model (I) the strategy is

different, as the system tends to favor regions with intermediate densities, typically ρ on the order of $1/2$. Of course, this is antagonistic to the trend of concentrating in regions where D is large and introduces some frustration in the system. At a very qualitative level, this is an explanation for the fact that multiplicity of the dynamically stable stationary state occurs in model (I) even when $d = 1$.

Acknowledgements

References

[Biler and Brandolese 2009] P. Biler and L. Brandolese, "On the parabolic-elliptic limit of the doubly parabolic Keller–Segel system modelling chemotaxis", *Studia Math.* **193**:3 (2009), 241–261.

[Biler et al. 2011] P. Biler, L. Corrias, and J. Dolbeault, "Large mass self-similar solutions of the parabolic-parabolic Keller–Segel model of chemotaxis", *J. Math. Biol.* **63**:1 (2011), 1–32.

[Blanchet and Laurençot 2013] A. Blanchet and P. Laurençot, "The parabolic-parabolic Keller–Segel system with critical diffusion as a gradient flow in $\mathbb{R}^d$, $d \geq 3$", *Comm. Partial Differential Equations* **38**:4 (2013), 658–686.

[Blanchet et al. 2006] A. Blanchet, J. Dolbeault, and B. Perthame, "Two-dimensional Keller–Segel model: optimal critical mass and qualitative properties of the solutions", *Electron. J. Differential Equations* (2006), Article ID #44.

[Blanchet et al. 2008] A. Blanchet, V. Calvez, and J. A. Carrillo, "Convergence of the mass-transport steepest descent scheme for the subcritical Patlak–Keller–Segel model", *SIAM J. Numer. Anal.* **46**:2 (2008), 691–721.

[Blanchet et al. 2015] A. Blanchet, J. A. Carrillo, D. Kinderlehrer, M. Kowalczyk, P. Laurençot, and S. Lisini, "A hybrid variational principle for the Keller–Segel system in $\mathbb{R}^2$", 2015. To appear in *ESAIM: M2AN*.

[Brezis 2011] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Springer, New York, 2011.

[Burger et al. 2006] M. Burger, M. Di Francesco, and Y. Dolak-Struss, "The Keller–Segel model for chemotaxis with prevention of overcrowding: linear vs. nonlinear diffusion", *SIAM J. Math. Anal.* **38**:4 (2006), 1288–1315.

[Burger et al. 2008] M. Burger, Y. Dolak-Struss, and C. Schmeiser, "Asymptotic analysis of an advection-dominated chemotaxis model in multiple spatial dimensions", *Commun. Math. Sci.* **6**:1 (2008), 1–28.

[Burger et al. 2010] M. Burger, M. Di Francesco, J.-F. Pietschmann, and B. Schlake, "Nonlinear cross-diffusion with size exclusion", *SIAM J. Math. Anal.* **42**:6 (2010), 2842–2871.

[Burger et al. 2011] M. Burger, P. A. Markowich, and J.-F. Pietschmann, "Continuous limit of a crowd motion and herding model: analysis and numerical simulations", *Kinet. Relat. Models* **4**:4 (2011), 1025–1047.

[Calvez and Carrillo 2012] V. Calvez and J. A. Carrillo, "Refined asymptotics for the subcritical Keller–Segel system and related functional inequalities", *Proc. Amer. Math. Soc.* **140**:10 (2012), 3515–3530.

[Calvez and Corrias 2008] V. Calvez and L. Corrias, "The parabolic-parabolic Keller–Segel model in $\mathbb{R}^2$", *Commun. Math. Sci.* **6**:2 (2008), 417–447.

[Di Francesco and Rosado 2008] M. Di Francesco and J. Rosado, "Fully parabolic Keller–Segel model for chemotaxis with prevention of overcrowding", *Nonlinearity* **21**:11 (2008), 2715–2730.

[Dolbeault et al. 2001] J. Dolbeault, P. A. Markowich, and A. Unterreiter, "On singular limits of mean-field equations", *Arch. Ration. Mech. Anal.* **158**:4 (2001), 319–351.

[Hörmander 1976] L. Hörmander, *Linear partial differential operators*, Grundlehren der Mathematischen Wissenschaften **116**, Springer, Berlin, 1976.

[Kirchner and Schadschneider 2002] A. Kirchner and A. Schadschneider, "Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics", *Physica A* **312**:1–2 (2002), 260–276.

[Laurençot and Matioc 2013] P. Laurençot and B.-V. Matioc, "A gradient flow approach to a thin film approximation of the Muskat problem", *Calc. Var. Partial Differential Equations* **47**:1-2 (2013), 319–341.

[Lopes 1996] O. Lopes, "Radial symmetry of minimizers for some translation and rotation invariant functionals", *J. Differential Equations* **124**:2 (1996), 378–388.

[Matthes et al. 2009] D. Matthes, R. J. McCann, and G. Savaré, "A family of nonlinear fourth order equations of gradient flow type", *Comm. Partial Differential Equations* **34**:10-12 (2009), 1352–1397.

JEAN DOLBEAULT: dolbeaul@ceremade.dauphine.fr
CEREMADE, CNRS UMR 7534, Université Paris-Dauphine, Place de Lattre de Tassigny, 75775 Paris Cedex 16, France

GASPARD JANKOWIAK: jankowiak@ceremade.dauphine.fr
CEREMADE, CNRS UMR 7534, Université Paris-Dauphine, Place de Lattre de Tassigny, 75775 Paris Cedex 16, France

PETER MARKOWICH: p.a.markowich@damtp.cam.ac.uk
Department of Applied Mathematics & Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom
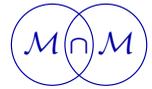
COMPREHENSIVE DESCRIPTION OF DEFORMATION OF SOLIDS AS WAVE DYNAMICS

SANICHIRO YOSHIDA

Deformation and fracture of solids are discussed as comprehensive dynamics based on a field theory. Applying the principle of local symmetry to the law of elasticity and using the Lagrangian formalism, this theory derives field equations that govern dynamics of all stages of deformation and fracture on the same theoretical foundation. Formulaically, these field equations are analogous to the Maxwell equations of electrodynamics, yielding wave solutions. Different stages of deformation are characterized by differences in the restoring mechanisms responsible for the oscillatory nature of the wave dynamics. Elastic deformation is characterized by normal restoring force generating longitudinal waves; plastic deformation is characterized by shear restoring force and normal energy-dissipative force generating transverse, decaying waves. Fracture is characterized by the final stage of plastic deformation where the solid has lost both restoring and energy-dissipative force mechanisms. In the transitional stage from the elastic regime to the plastic regime where both restoring and energy-dissipative normal force mechanisms are active, the wave can take the form of a solitary wave. Experimental observations of transverse, decaying waves and solitary waves are presented and discussed based on the field theory.

1. Introduction

Most conventional approaches classify deformation of solids into separate regimes and discuss its mechanics based on the constitutive relation of each regime. The elastic and plastic regimes are defined, respectively, as the regimes where the constitutive relation is characterized by linear and nonlinear stress-strain relations, and fracture is considered as an independent phenomenon where a preexisting crack grows by itself. For each regime, a specific theory is used: continuum mechanics [Spencer 1980; Landau and Lifshitz 1986] for the elastic regime, various theories of plasticity [Hill 1998; Lubliner 2008] for the plastic regime and fracture mechanics [Griffith 1921; Irwin 1948] for the fracturing regime. When an external load is applied to a solid object and the resultant deformation is viewed as the response

of the entire object, it is true that the deformation mechanics undergone in these regimes exhibits the stress-strain characteristics of the respective regimes. If the same object is analyzed in a local region, however, it is obvious that the deformation status can be different from other local regions, and hence, on the global level, the deformation should be characterized by multiple constitutive relations at the same time. A metal specimen freshly taken out of an annealing oven has a number of dislocations; as soon as an external load is applied to such a specimen, the dislocations are activated causing local plastic deformation. These local nonlinear behaviors are not observed in the stress-strain characteristic of the entire specimen as most of the specimen undergoes purely elastic deformation at this stage. At the other extreme, a metal specimen about to fail recovers from the deformed state if the load is removed. It is unrealistic to rely on a regime-specific theory in any stage of deformation. It is essential to use a theory that can describe all stages of deformation comprehensively.

Comprehensive theory of deformation and fracture is not only useful to describe the situation where elastic and plastic deformations coexist. It is also essential to formulate the transitions from one regime to another. Generally, the life of solids under external loads is a progression from elastic deformation to fracture through plastic deformation. In the tensile or compression mode of deformation where the stress increases with the passage of time, the deformation exhibits this pattern of progression as a function of the increasing stress. Even in the fatigue mode of deformation where the magnitude of the external load remains the same, most solids follow the same pattern of progression [Ichinose et al. 2006]. In engineering, often analysis of the transitional stage from one regime to the next is more important than analysis within a certain regime. If the remaining life of a machine part is known, it becomes unnecessary to replace it at an earlier stage, contributing to the conservation of natural resources. To analyze these transitional stages, the theory must be regime-independent. Furthermore, these transitions involve multiple scale levels. Fracture of solids is always initiated at the atomistic scale and evolves to the microscopic scale and eventually to the macroscopic scale; defects of a size comparable to several atoms grow to the microscopic level and eventually to the macroscopic level when the entire specimen fails. A universal approach is essential.

In this regard, a recent field theory of deformation and fracture has an advantage [Yoshida 2015]. Requesting local symmetry [Aitchison and Hey 1989] in Hooke's law, this theory derives field equations that govern the displacement field of solids under deformation. Formulaically, the field equations are very similar to Maxwell's equations of electrodynamics. From a dynamical point of view, the field equations represent synergetic interaction between the translational and rotational modes of displacement. This interaction can be interpreted as Lenz's law

analogous to Faraday's law in electrodynamics. In the present context, Lenz's law represents solids' response to reduce the disturbance caused by an external load. The general solution to the field equations is a wave function where the form of the wave characteristics depends on the regime. The elastic regime can be characterized by a longitudinal wave known as the wave of compression and a rotational wave known as the wave of deformation. In this regime, the elasticity is a longitudinal effect where the solid material responds to the force due to an external load elastically. The plastic regime is characterized by a transverse wave that decays due to the irreversibility of plasticity. In this regime, the elasticity is a transverse effect due to differential rotation; the solid material responds elastically to the external torque and not to the force. The irreversibility is due to the energy dissipation associated with irreversible motion of localized normal strain[1] in the direction of the local velocity. Thus, the longitudinal effect is energy-dissipative. The solid material resists the external force energy-dissipatively and the external torque elastically. Fracture is the stage of deformation where the solid material has lost all the mechanisms to resist the external load, elastically or energy-dissipatively, and the only reaction to the load is to create discontinuity. Under some conditions, the energy-carrying wave takes the form of a solitary wave. In this situation, the solid material does not exhibit resistive force, and the stress-strain curve plateaus. A similar phenomenon occurs in the transitional stage from the plastic to fracture regimes. The material dissipates energy from the external load via propagation of a solitary wave. When the solitary wave stops moving, the material loses the energy-dissipative mechanism completely, and it fractures. Thus, the transition from one regime of deformation to another can be identified as a change in the way the solid material responds to an external load and characterized as different forms in the displacement wave.

The similarity between electrodynamics and the present field deformation theory is not limited to the formulaic resemblance in the field equations. As mentioned above, the energy dissipation in plastic deformation is caused by motion of localized normal strain due to the local velocity field. From the field-theoretical viewpoint, the normal strain can be interpreted as the charge of symmetry associated with the local symmetry of Hooke's law. From the viewpoint of Lagrangian dynamics, the normal strain can be interpreted as representing the Noether current associated with the invariance of the corresponding Lagrangian density. From this standpoint, this quantity can be identified as corresponding to the electric charge and called the deformation charge. Note that the electric charge is proportional to divergence of the electric field and the normal strain is divergence of the displacement field. With this interpretation, the energy dissipation in plastic deformation

---

[1]Strictly speaking, it is the rate of normal strain or the temporal derivative of normal strain. For simplicity, it is called the normal strain.

can be understood as a phenomenon analogous to Ohmic loss in electrodynamics where the electromagnetic field loses energy when a free electric charge is moved by the electric field. From the perspective of energy flow, the transverse displacement wave carries elastic energy through the material and the charge flow dissipates the energy. When the transverse wave decays completely and the charge stops moving, the material loses all mechanisms to transfer the work provided by the external load from one side of the specimen to the other or dissipate it. This is when the material fractures. The situation is analogous to electrical breakdown of dielectric media [Yoshida 2000]. When the level of ionization is low and hence the density of free charge is low, the propagation of electromagnetic waves is the dominant mechanism of energy flow. When the conduction current density increases to a certain level, the Ohmic loss becomes the dominant mechanism to dissipate the energy provided by the external circuit. Eventually, the current density becomes infinitely high, and that is when the medium is electrically broken.

A number of authors have formulated nonlinear behavior of deformation. Among them, the following models have been proposed as useful tools for unified description of elastic and plastic deformations and are worth mentioning here. These models are based on the framework established by Toupin [1964] and Mindlin [1965] and known as Toupin–Mindlin strain-gradient theory. This theory postulates that the strain energy depends both on the symmetric strain tensor and the second gradient of displacement. By introducing a Lagrangian action both in the material and the spatial description, Auffray et al. [2015] have formulated a material description for second-gradient continua. Javili et al. [2013] have generalized the work by Mindlin [1965] and formulated a geometrically nonlinear theory of higher-gradient elasticity accounting for boundary energies. By means of the least action principle, Madeo et al. [2013] have derived a general set of equations of motion and duality conditions to be imposed at macroscopic surfaces of discontinuity in partially saturated, solid second-gradient porous media. Fleck and Hutchinson [1997] have applied the Toupin–Mindlin strain-gradient framework to plastic deformation and proposed phenomenological theories of strain-gradient plasticity.

Connections between these formalisms based on the strain-gradient theory and the present field theory are not straightforward and are not fully understood at this point. Nevertheless, it is worth pointing out some similarities and a contrast between these formalisms and the present theory. As will be discussed in the next section, the present theory incorporates nonlinearity by allowing the transformation matrix of linear elasticity, known as the displacement gradient tensor, to be coordinate-dependent. The components of the displacement gradient tensor are strain, which is essentially the first gradient of displacement. Thus, the fact that we allow its coordinate dependence means that we automatically consider the second gradient of displacement. Naively speaking, this corresponds to the

approach taken by the strain-gradient theory where the second gradient of dis-placement is included in the expression of the strain energy. Auffray et al. [2015] and Madeo et al. [2013] apply the Lagrangian formalism to deduce the evolution equations. This procedure seems analogous to the present theory whereby the field equations are derived through the application of the Lagrangian formalism to the gauge potential. Moreover, the use of both the material and spatial descriptions in the Lagrangian action made by Auffray et al. indicates that the two descrip-tions interact with each other in the same or similar fashion as the gauge field and the linear elastic displacement field interact with each other in the present theory.

As for the contrast, the following point should be noted. In the present field theory, the nonlinearity is introduced in conjunction with the coordinate depen-dence of the transformation matrix that describes the local linear elasticity (the base theory). In other words, the nonlinearity is associated with the curvilinearity of the coordinate axes and not intrinsic in the base theory. In the case of the strain-gradient theory, on the other hand, the starting Lagrangian encompasses nonlinearity as the strain energy expression has the term containing the second gradient of displacement; thus, the formalism derived from the strain-gradient the-ory is intrinsically able to describe the nonlinear nature of geometry such as the inclusion of porosity and layered structures. It may be possible to incorporate these geometrical nonlinear effects by the use of an appropriate compensation field, but the possibility is not clear at this time. It is safe to assume, at least for now, that the present theory is applicable to the cases where the local deformation can be modeled to obey the law of linear elasticity.

The aim of this paper is to provide an overview of this field-theoretical approach to dynamics of deformation and fracture. After briefly reviewing the concept of local symmetry on which the present theory is based, the field equations will be derived. Through physical interpretations of the field equations, deformation dy-namics will be discussed from the viewpoint of force acting on a unit volume of the solid material. It will be shown that the transverse (shear) force is restoring regardless of the regime whereas the longitudinal (normal) force is restoring in the elastic regime and energy-dissipative in the plastic regime. It will also be shown that these restoring forces cause longitudinal and transverse wave nature of elastic and plastic deformation and solitary wave nature in the transitional stage from the elastic regime to the plastic regime. The wave equations for the respective regimes will be derived from the field equations and will be solved analytically under some conditions. The energy-dissipative nature of the dynamics will be explained through the concept of deformation charge. Supporting experimental results will be presented to discuss these wave characteristics and dynamics of the deformation charge.

2. Theory

Details of the present field theory can be found elsewhere [Yoshida 2015]. In short, this field theory formulates deformation dynamics based on two postulates. The first postulate is that a solid of any deformation status locally obeys the law of linear elasticity (Hooke's law). The local region that obeys Hooke's law is referred to as the deformation structural element. The second postulate is that, as long as the solid remains a continuum, all the deformation structural elements of the object are logically connected by a field known as the gauge field. The first postulate is rationalized through the consideration that regardless of the stage of deformation it is always possible to find a local region where the interatomic potential is approximated by a quadratic function of the displacement of the atom from its equilibrium position or, equivalently speaking, the field force on the atom is elastic force whose magnitude is proportional to the displacement. The above-mentioned claim that a solid object about to fail recovers from the deformed state to a certain extent if the load is removed is an example. This postulate indicates that each of these local elastic deformations can be expressed by a transformation that represents Hooke's law with the local coordinate system (local frame). If the local frame is oriented to the principal axis, the corresponding transformation matrix is diagonal as the shear components of the stress and strain matrices are all zero. Here, it is important to note that the principal axes of the deformation structural elements within the same object do not necessarily have the same orientation. In fact, it is usually the case that in the plastic regime they are oriented randomly, as will be discussed later. This means that we cannot define a principal axis with the global coordinate system (global frame) and that therefore we cannot express the local elastic deformations inclusively. This is where the second postulate comes into the picture.

The second postulate can be argued in various ways based on the principle of local symmetry. The most intuitive argument will be as follows. The situation where multiple deformation structural elements undergo linear elastic deformation with the respective principal axes raises a question: "are the local elastic deformations expressed in the respective local coordinates independent of one another?" or, equivalently, "do the deformation structural elements know one another's elastic deformation?" The answer must be: "they are not independent of one another" or "they should know one another's behavior". Otherwise, the situation becomes the same as the same number of independent solid objects (not connected with one another) as the deformation structural elements experience elastic deformation independently. Then the next question will be: "how are they connected?" We can find a short answer to this question by recalling that we try to express all the local elastic deformations inclusively with the global coordinates under the situation where the local elastic deformations have their own orientations of the
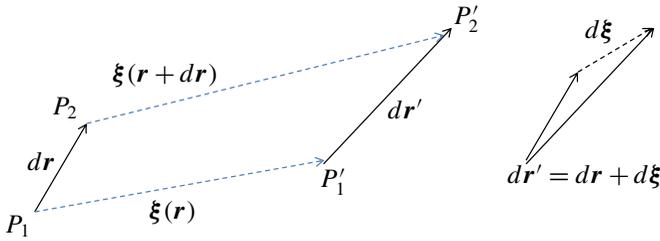
Figure 1. Infinitesimal line element changes its length under deformation.

principal axis. In other words, each deformation structural element experiences stretch and compression along mutually different orientations. In order to express these behaviors inclusively in the global frame, it is necessary to align all these axes in the same orientation. This argument indicates that the gauge field has something to do with rotational dynamics. As will be discussed later in this paper, indeed the present gauge field is associated with rotational dynamics. From the field-theoretical viewpoint, the gauge field compensates for the fact that the elastic deformation expressed in the global coordinate does not consider the nonlinearity of the dynamics. From this standpoint, this field is referred to as the compensation field. The portion of the dynamics that we overlook by artificially aligning the deformation structural elements, which is responsible for the nonlinearity and the irreversibility of plastic deformation, is crammed into the potential generated by the gauge field. Naturally, this potential is rotation-like. We say that the gauge field makes the law of linear elasticity locally symmetric. Mathematically, it can be stated as follows. Under the situation where deformation structural elements undergo respective elastic deformation, the transformation matrix is coordinate-dependent. Consequently, the associated physics law cannot be written with the global coordinates in the same form as the local coordinates. This is because the expression of the physics law involves differentiation and the coordinate dependence of the transformation matrix generates the extra term resulting from the differentiation of the matrix. The gauge field regains the formality in the global frame by adding an extra term (the gauge term) to the usual derivatives so that this term cancels out the effect of the differentiation of the transformation matrix. The derivative with the gauge term is referred to as the covariant derivative. It is interesting to note that mathematicians call the gauge field the connection field. In the present case, the gauge field literally connects deformation structural elements so that they form a single continuum.

2.1. *Deformation as linear transformation.* Consider in Figure 1 a solid object under deformation. By this deformation, an infinitesimal line element vector $\boldsymbol{\eta} = (dx, dy, dz)$ changes its length and direction by coordinate-dependent displacement vector $\boldsymbol{\xi}(x, y, z)$. Expressing the resultant line element vector as $\boldsymbol{\eta}'$, we can

express the deformation with the linear transformation

$$\boldsymbol{\eta}' = (I + \beta) \equiv U\boldsymbol{\eta}, \tag{1}$$

where I is the unit matrix and β is the displacement gradient tensor

$$\beta = \left(\delta_{ij} + \frac{\partial \xi_i}{\partial x^j} \right). \tag{2}$$

Here, $\boldsymbol{\xi}$ is the displacement vector. In the theory of elasticity, the elastic force is proportional to the stretch or the differential displacement $d\boldsymbol{\xi}$. Therefore, for the theory to be invariant, the differential displacement vector must be transformed in the same fashion as the displacement vector itself. Otherwise, after the transformation, the elastic force law cannot be written in the same form as before the transformation. This means that, after the transformation, the differential displacement must have the same form as the transformation of the differential displacement. If we use the usual differentiation, apparently this is not the case as the following expression indicates:

$$d(U\boldsymbol{\xi}) = dU\boldsymbol{\xi} + U d\boldsymbol{\xi}. \tag{3}$$

The condition $d(U\boldsymbol{\xi}) = U d\boldsymbol{\xi}$ holds only when $dU = 0$ or the transformation is coordinate-independent. Thus, it becomes necessary to replace the usual derivatives with covariant derivatives or introduce a gauge term Γ_i:

$$D_i = \frac{\partial}{\partial x_i} - \Gamma_i \equiv \partial_i - \Gamma_i. \tag{4}$$

It is easily proved that, if the gauge term transforms as (5), the differential after the transformation has the same form as the transformation of the differential, that is, $D_i'(U\boldsymbol{\xi}) = U(D_i\boldsymbol{\xi})$:

$$\Gamma_i' = U\Gamma_i U^{-1} + \frac{\partial U}{\partial x^i} U^{-1}. \tag{5}$$

Here the prime $'$ indicates the quantity after the transformation.

Now consider the physical meaning of the gauge term:

$$D\xi_i = \left(\frac{\partial \xi_i}{\partial x} - \Gamma_x \xi_i \right) dx + \left(\frac{\partial \xi_i}{\partial y} - \Gamma_y \xi_i \right) dy + \left(\frac{\partial \xi_i}{\partial z} - \Gamma_z \xi_i \right) dz \equiv d\xi_i - A_i. \tag{6}$$

In elastic deformation, the rotation matrix represents rigid body rotation of the material, which does not involve length change. In (6), the actual change in the length of the displacement vector is all in $d\xi_i$. Thus, A_i can be interpreted as representing a displacement vector that rotates the deformation structural element so that the differential displacement vector contains only the change due to physically true deformation and not to the geometrical effect. Figure 2 illustrates this situation
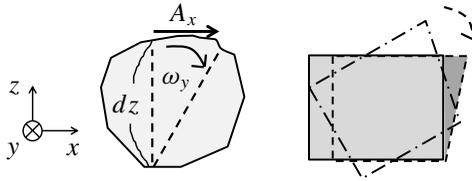
Figure 2. Vector potential as displacement vector to align deformation structural element.

schematically. Here ω_y represents the rotation that must be removed to find out the true deformation.

The above argument that the vector potential represents material rotation associated with the covariant derivatives can be justified from the following viewpoint. In the theory of elasticity, differential displacement (deformation gradient tensor) can be separated into the symmetric and asymmetric portions referred to as the strain and rotation matrices, respectively. When the coordinate axes are chosen to be the principal axes, the strain matrix is diagonal, or its shear components are all zero. Now consider that different parts of a given solid object undergo their respective elastic deformations. As an example, when an initially elastic object has a defect, the four blocks will undergo different elastic deformation as Figure 3 illustrates schematically. Under this condition, the four blocks have their own principal axes. It becomes impossible to describe the four elastic deformations with a common principal axis.

Based on the above interpretation, the explicit form of the spatial parts of vector potential A can be identified as

$$A = \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} -\omega_z dy + \omega_y dz \\ \omega_z dx - \omega_x dz \\ -\omega_y dx + \omega_x dy \end{pmatrix}. \tag{7}$$

The temporal component of vector potential A can be understood in conjunction with the temporal differentiation as follows. Suppose deformation dynamics ψ
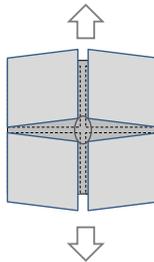
Figure 3. Local region containing a defect at the center.
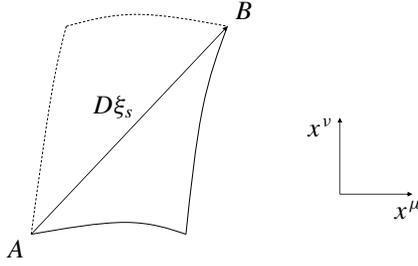
Figure 4. Gauge field strength for nonlinear deformation.

propagates as a wave. In general, the wave function can be put in the form

$$\psi = f(\omega_0 t - \boldsymbol{k} \cdot \boldsymbol{r}) = f(\omega_0 t - k(\alpha x + \beta y + \gamma z)), \tag{8}$$

where ω_0 is the angular frequency, $\boldsymbol{k}$ is the propagation vector and α, β and γ are directional cosines. Expressing the derivative of function f as f', we find

$$f' = \frac{1}{\omega_0} \frac{\partial \psi}{\partial t} = -\frac{1}{k}(\nabla \psi) \cdot \hat{k}, \tag{9}$$

where $\hat{k}$ is the unit vector of $\boldsymbol{k}$. Noting the ratio of the angular frequency to the propagation vector is the phase velocity c, $\omega_0/k = c$, we can rewrite (9) as[2]

$$\frac{\partial \psi}{\partial t} = -(\nabla \psi) \cdot c\hat{k} = -(\nabla \psi) \cdot \boldsymbol{c}. \tag{10}$$

The spatial and temporal components of vector potential $\boldsymbol{A}$ are to compensate the spatial and temporal differentiations. Thus, they can be interpreted as being corresponding to terms $(\nabla \psi)$ and $\partial \psi / \partial t$ in (10). This interpretation leads to the vector potential expression of $\boldsymbol{A}$ as follows. Intuitively, the temporal component of the vector potential can be interpreted as representing the same effect as the spatial component explained in Figure 2; it represents the effect of the compensating potential in the time domain wherein the dynamics is a wave phenomenon traveling at the phase velocity c:

$$A^\mu = (A^0, A^1, A^2, A^3) = \left(\frac{\phi^0}{c}, A^1, A^2, A^3\right). \tag{11}$$

Now consider how vector potential $\boldsymbol{A}$ acts on the deformation dynamics. We know that $\boldsymbol{A}$ represents the dynamics that local linear elastic dynamics cannot represent. In other words, it accounts for the compensation we need to pay as the penalty for pretending that the dynamics is linear elastic. This effect can be formulated by

---

[2]If we repeat the same procedure to find the secondary derivatives, we will obtain the elastic wave equation where the phase velocity is the square root of the ratio of the elastic modulus to the density.

comparing clockwise and counterclockwise covariant derivatives. Figure 4 illustrates the clockwise and counterclockwise differentiation schematically. Dropping the second-order differentials, the clockwise case is

$$D_\mu(D_\nu\xi_s dx^\nu)dx^\mu = \partial_\mu\partial_\nu\xi_s dx^\nu dx^\mu - \partial_\mu(\Gamma_\nu\xi_s d^\nu)dx^\mu + \Gamma_\mu\Gamma_\nu\xi_s dx^\nu dx^\mu. \quad (12)$$

Here from the definition, $\Gamma_\nu\xi_s dx^\nu$ can be interpreted as A_ν, and so

$$D_\mu(D_\nu\xi_s dx^\nu)dx^\mu = \partial_\mu\partial_\nu\xi_s dx^\nu dx^\mu - \partial_\mu A_\nu dx^\mu + \frac{1}{\xi_s}A_\mu A_\nu. \quad (13)$$

The counterclockwise case can be expressed with the vector potential in the same fashion. Thus, the difference between the clockwise and counterclockwise cases is

$$D_\mu(D_\nu\xi_s dx^\nu)dx^\mu - D_\nu(D_\mu\xi_s dx^\mu)dx^\nu$$
$$= (\partial_\nu A_\mu dx^\nu - \partial_\mu A_\nu dx^\mu) + \frac{1}{\xi_s}[A_\mu, A_\nu]. \quad (14)$$

In the infinitesimal limit, $dx^\nu = dx^\mu = ds$, and division of the above equation by ds leads to

$$[D_\mu, D]_s\xi_s ds = (\partial_\nu A_\mu - \partial_\mu A_\nu) + \frac{1}{\xi_s ds}[A_\mu, A_\nu] \equiv F_{\mu\nu}. \quad (15)$$

Here $F_{\mu\nu}$ is known as the field stress tensor. Each component of vector potential equation (11) represents a displacement component. It is easily proved that they are commutable; hence, the $[A_\mu, A_\nu]$ term of (15) is zero. With this, we obtain the explicit form

$$F_{\mu\nu} = \begin{pmatrix} 0 & -v^1/c & -v^2/c & -v^3/c \\ v^1/c & 0 & -\omega^3 & \omega^2 \\ v^2/c & \omega^3 & 0 & -\omega^1 \\ v^3/c & -\omega^2 & \omega^1 & 0 \end{pmatrix}. \quad (16)$$

Here v^i, $i = 1, 2, 3$, is the time derivative of A^i and ω^i, $i = 1, 2, 3$, is the rotation associated with the corresponding components of the displacement due to vector potential A:

$$\omega^k = \frac{\partial A^j}{\partial x^i} - \frac{\partial A^i}{\partial x^j}. \quad (17)$$

2.2. Lagrangian and field equation. The field stress tensor is not invariant under the transformation U. However, it is easily proved that the trace $F_{\mu\nu}F^{\mu\nu}$ is invariant [Yoshida 2011]. This indicates that we can construct the Lagrangian of free particles (without the interaction with the gauge field or vector potential) in the form proportional to $F_{\mu\nu}F^{\mu\nu}$. Knowing that the phase velocity c_{shear} associated with shear force has the form

$$c_{\text{shear}} = \sqrt{G/\rho}, \quad (18)$$

where G is the shear modulus and ρ is the density, and that the Lagrangian is kinetic energy minus potential energy, we can identify the Lagrangian density as

$$L = -\frac{G}{4}F_{\mu\nu}F^{\mu\nu} = \frac{\rho v^2}{2} - \frac{G\omega^2}{2}. \tag{19}$$

Here the first term is the kinetic energy of the unit volume and the second term is the rotational spring potential energy. The Lagrangian in the form of (19) indicates the phase velocity c in (16) is in fact the shear wave velocity (18). By adding the interaction term, we can identify the full Lagrangian in the form

$$L = -\frac{G}{4}F_{\mu\nu}F^{\mu\nu} + Gj^\mu A_\mu = \frac{\rho v^2}{2} - \frac{G\omega^2}{2} + \frac{G}{c}j^0 A_0 + Gj^i A_i. \tag{20}$$

Here j^0 and j^i are the temporal and spatial components of the quantity known as the charge of symmetry, and they are connected with the phase velocity (18) as

$$j^\mu = \left(\frac{j^0}{c}, j^1, j^2, j^3\right). \tag{21}$$

The four vector j^μ describes how the material interacts with the gauge field and is conserved under the governing transformation (in this case transformation U).

With this Lagrangian, the Euler–Lagrangian equation of motion associated with A_μ can be given as

$$\partial_\nu \frac{\partial L}{\partial(\partial_\nu A_\mu)} - \frac{\partial L}{\partial A_\mu} = 0. \tag{22}$$

This leads to the following field equations:

$$\nabla \cdot \boldsymbol{v} = -j_0, \tag{23}$$

$$\nabla \times \boldsymbol{v} = \frac{\partial \boldsymbol{\omega}}{\partial t}, \tag{24}$$

$$\nabla \times \boldsymbol{\omega} = -\frac{1}{c^2}\frac{\partial \boldsymbol{v}}{\partial t} - \boldsymbol{j}, \tag{25}$$

$$\nabla \cdot \boldsymbol{\omega} = 0. \tag{26}$$

Here c appearing on the right-hand side of (25) is the phase velocity (18)

2.3. *Deformation charge and comprehensive description.* Rearranging the terms, we can put the field equation (25) in the form [Yoshida 2011; 2008]

$$\frac{1}{c^2}\frac{\partial \boldsymbol{v}}{\partial t} = -\nabla \times \boldsymbol{\omega} - \boldsymbol{j}. \tag{27}$$

The spatial component of the charge j^μ appearing on the right-hand side of the third field equation (25) represents the longitudinal effect of the gauge field on the material and is very important to describe the dynamics. The c appearing in the first
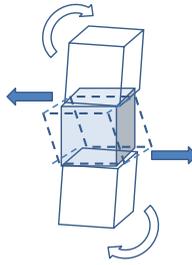
Figure 5. Shear force due to differential rotational displacement.

term on the right-hand side of the same equation represents the phase velocity (18), the velocity that the spatial pattern of the displacement field propagates. Using the phase velocity expression (18) and multiplying both sides of this equation by ρ, we can put (27) in the form

$$\rho \frac{\partial \boldsymbol{v}}{\partial t} = -G\nabla \times \boldsymbol{\omega} - G\boldsymbol{j}. \tag{28}$$

The left-hand side of (28) is found to have the form of the product of the mass and acceleration of the unit volume. Thus, according to Newton's second law, the right-hand side of (27) is the external force acting on the unit volume. Here the first term $G\nabla \times \boldsymbol{\omega}$ represents the shear force exerted by the neighboring blocks of the material due to their differential rotations, and the second term $G\boldsymbol{j}$ can be identified as the longitudinal force density. Figure 5 illustrates the shear force schematically. The form of this second term differentiates different regimes of deformation, as will be discussed below.

Elastic regime. Take the divergence of both sides of (25) and substitute the resultant $\nabla \cdot \boldsymbol{v}$ with (23). This provides us with an equation of continuity associated with the conservation of charge $j_0 = -\nabla \cdot \boldsymbol{v}$:

$$\rho \frac{\partial (\nabla \cdot \boldsymbol{v})}{\partial t} = -\nabla \cdot (G\boldsymbol{j}). \tag{29}$$

Using $\partial \boldsymbol{\xi} / \partial t = \boldsymbol{v}$, rewrite the left-hand side of (29) as

$$\rho \frac{\partial^2 (\nabla \cdot \boldsymbol{\xi})}{\partial t^2} = -\nabla \cdot (G\boldsymbol{j}). \tag{30}$$

The quantity $\nabla \cdot \boldsymbol{\xi}$ appearing on the left-hand side of (30) is known as the volume expansion in continuum mechanics. With the interpretation that $G\boldsymbol{j}$ represents the longitudinal force, (30) can be interpreted as the equation of motion of a unit volume experiencing volume expansion due to differential longitudinal (normal) force. We can identify the explicit form of $G\boldsymbol{j}$ for an isotropic elastic medium based on continuum mechanics.

Recall that the constitutive relation can be written in the following form based on Cauchy's formalism:

$$
\begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{xy} \\ \sigma_{yz} \\ \sigma_{zx} \end{pmatrix} = \begin{pmatrix} \lambda+2G & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda+2G & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda+2G & 0 & 0 & 0 \\ 0 & 0 & 0 & 2G & 0 & 0 \\ 0 & 0 & 0 & 0 & 2G & 0 \\ 0 & 0 & 0 & 0 & 0 & 2G \end{pmatrix} \begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ \epsilon_{xy} \\ \epsilon_{yz} \\ \epsilon_{zx} \end{pmatrix}, \tag{31}
$$

where σ_{ij} denotes the j-th component of the stress acting on plane i, λ is the first Lamé coefficient and ϵ_{ij} is the strain defined as

$$
\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial \xi_j}{\partial x_i} + \frac{\partial \xi_i}{\partial x_j}\right). \tag{32}
$$

Considering the x component of the net external force acting on a cube of unit volume, we obtain the equation of motion

$$
\rho\frac{\partial^2 \xi_x}{\partial t^2} = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{yx}}{\partial y} + \frac{\partial \sigma_{zx}}{\partial z}. \tag{33}
$$

Substituting the corresponding stress tensor components of the constitutive equation (31) into the right-hand side and using (32), we can rewrite

$$
\rho\frac{\partial^2 \xi_x}{\partial t^2} = G\nabla^2 \xi_x + \frac{\partial}{\partial x}(\lambda+G)\nabla \cdot \boldsymbol{\xi}. \tag{34}
$$

View $G\nabla^2 \xi_x$ as the x component of $G\nabla^2 \boldsymbol{\xi}$, and use the mathematical identity $\nabla \times \nabla \times = \nabla(\nabla \cdot) - \nabla^2$ to rewrite $G\nabla^2 \boldsymbol{\xi}$ as

$$
G\nabla^2 \boldsymbol{\xi} = -G\nabla \times \nabla \times \boldsymbol{\xi} + G\nabla(\nabla \cdot \boldsymbol{\xi}). \tag{35}
$$

On the right-hand side of (35), the longitudinal effect is represented by the second term. Taking only this term and noting that the x component of $\nabla(\nabla \cdot \boldsymbol{\xi})$ can be put as $(\nabla(\nabla \cdot \boldsymbol{\xi}))_x = \partial(\nabla \cdot \boldsymbol{\xi})/\partial x$, we can rewrite (34) as

$$
\rho\frac{\partial^2 \xi_x}{\partial t^2} = G(\nabla(\nabla \cdot \boldsymbol{\xi}))_x + \frac{\partial}{\partial x}(\lambda+G)(\nabla \cdot \boldsymbol{\xi}) = \frac{\partial}{\partial x}(\lambda+2G)(\nabla \cdot \boldsymbol{\xi}). \tag{36}
$$

Including the y and z components, we can express the longitudinal force vector appearing on the right-hand side of (28) in the form

$$
G\boldsymbol{j} = -\nabla(\lambda+2G)(\nabla \cdot \boldsymbol{\xi}). \tag{37}
$$

We can identify this form of $G\boldsymbol{j}$ as the longitudinal force in the elastic case. Here $(\lambda+2G)(\nabla \cdot \boldsymbol{\xi})$ is the longitudinal force at a surface of the unit-volume cube proportional to the strain at that point, and the entire right-hand side represents
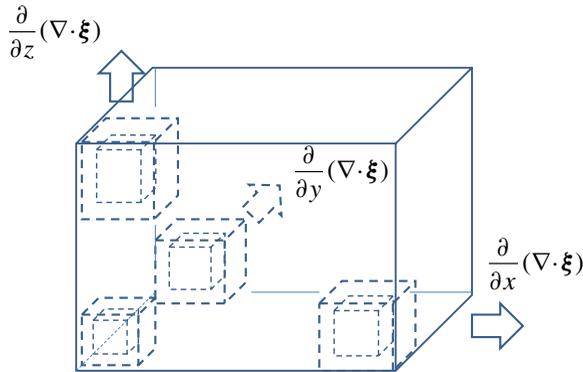
Figure 6. Elastic force proportional to volume expansion.

the differential longitudinal force, the difference of $(\lambda + 2G)(\nabla \cdot \boldsymbol{\xi})$ between the leading and tailing surfaces of the unit-volume cube. Figure 6 illustrates the force schematically.

Note that the first term on the right-hand side of (35) represents the shear force and can be written as $\nabla \times \nabla \times \boldsymbol{\xi} = \nabla \times \boldsymbol{\omega}$. This terms appears as the shear force term in (28). In the elastic limit, (35), hence field equation (25), reduces to Cauchy's constitutive equation.

From the viewpoint of the equation of continuity, (29) indicates that compression or rarefaction of en elastic material does not appear or disappear by itself but is only generated by longitudinal force exerted by the neighboring volume. It is interesting to note that compression and rarefaction occur alternatively.

Plastic regime. Viewing (29) as an equation of continuity, Gj can be interpreted as a flow of charge $\rho\nabla \cdot \boldsymbol{v}$. Thus, we can put

$$Gj = W_d \rho \nabla \cdot \boldsymbol{v}. \tag{38}$$

Here W_d is the drift velocity of the charge $\rho\nabla \cdot \boldsymbol{v}$ of the unit volume. As will be discussed below, if the charge is positive, it flows in the same direction as the local velocity of the material to dissipate the kinetic energy carried by the material particles. If it is negative, the charge flows in the direction opposite to the local velocity. Thus, we can put

$$W_d = \sigma_0 \boldsymbol{v}. \tag{39}$$

Here σ_0 is a dimensionless parameter that represents the degree of energy dissipation; the greater σ_0 is, the more energy is dissipated. With (38) and (39), (28) can be put in the form

$$\rho\frac{\partial \boldsymbol{v}}{\partial t} = -G\nabla \times \boldsymbol{\omega} - \sigma_0\rho(\nabla \cdot \boldsymbol{v})\boldsymbol{v}$$

$$= -G\nabla \times \boldsymbol{\omega} - \sigma_c\boldsymbol{v}. \tag{40}$$
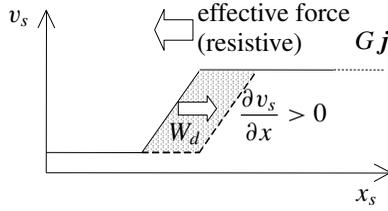
Figure 7. Positive charge flowing in the direction of local particle velocity.

The first term on the right-hand side of (40) represents recovery force due to shear deformation. Being proportional to the velocity, the second term can be interpreted as representing a velocity damping force, where

$$\sigma_c = \sigma_0 \rho (\nabla \cdot \boldsymbol{v}) \tag{41}$$

is the damping coefficient. This effect is interpreted as the energy-dissipative nature of plastic deformation.

Consider the physical meaning of the damping coefficient σ_c and the dimensionless parameter that represents the degree of energy dissipation σ_0. To this end, rewrite the $\nabla \cdot \boldsymbol{v}$ that appears on the right-hand side of (41) as

$$\begin{aligned}
\nabla \cdot \boldsymbol{v} &= \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \\
&= \frac{\partial v_s}{\partial x}\alpha + \frac{\partial v_s}{\partial y}\beta + \frac{\partial v_s}{\partial z}\gamma \\
&= \frac{\partial v_s}{\partial x}\frac{dx}{dx_s} + \frac{\partial v_s}{\partial y}\frac{dy}{dx_s} + \frac{\partial v_s}{\partial z}\frac{dz}{dx_s} = \frac{dv_s}{dx_s}.
\end{aligned} \tag{42}$$

Here α, β and γ are direction cosines and $\boldsymbol{v}_s$ is the local velocity vector of the particle. With (39) and (42), we can put the plastic longitudinal force density (38) in the form

$$G\boldsymbol{j} = \sigma_0 \rho \frac{dv_s}{dx_s} v_s \hat{v}_s. \tag{43}$$

Consider the right-hand side of (43) for the case $\sigma_0 = 1$. Since $v_s = dx_s/dt$, it can be put as

$$G\boldsymbol{j} = \sigma_0 \rho \frac{dv_s}{dx_s} v_s \hat{v}_s = \rho \frac{dv_s}{dx_s}\frac{dx_s}{dt}\hat{v}_s = \rho \frac{dv_s}{dt}\hat{v}_s = \frac{d}{dt}(\rho v_s)\hat{v}_s. \tag{44}$$

The rightmost side of (44) represents the temporal change of momentum density (the momentum change of the unit volume). Thus, (43) can be interpreted as indicating the effect that the external longitudinal force has to cause the unit volume to gain momentum. In other words, when $\sigma_0 = 1$, the external force exerted by the neighboring volume contributes to the momentum increase of this unit volume without energy dissipation. From (38), we know that the charge drifts in the same
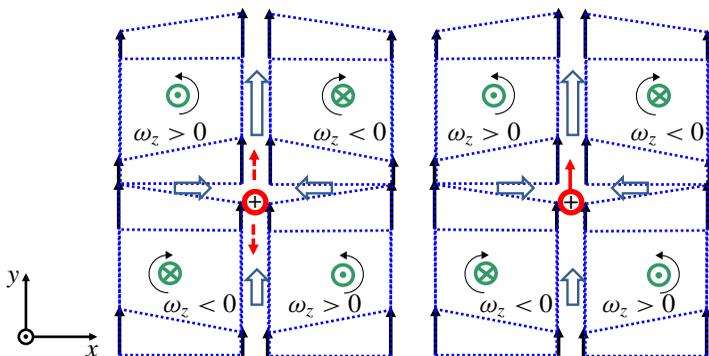
Figure 8. Schematic illustration of shear strain. Four blocks of the specimen under tensile load experience rotation-like behaviors due to differential vertical displacement indicated with narrow arrows at corners of blocks. Thicker arrows indicate resulting displacement along boundaries of the four blocks. The blocks' rotations generate tensile strain near the center of the specimen. Left: the elastic situation where the volumes neighboring this central volume at the top and bottom resist the tensile strain by exerting elastic force that can be viewed as the shear force represented by $\nabla \times \boldsymbol{\omega}$. Right: the plastic situation where the material yields to the shear force and consequently the central volume drifts upward. This can be viewed as a positive charge drifts upward causing energy dissipation due to the friction exerted by surrounding materials.

direction as the longitudinal force at the drift velocity $\boldsymbol{W}_d$. From (43) when $\sigma_0 = 1$, $\boldsymbol{W}_d$ is equal to the particle velocity $\boldsymbol{v}$. So in this case, the charge dv_s/dx_s drifts in the direction of the longitudinal force at the particle velocity v_s.

When $\sigma_0 > 1$, the situation is slightly different. In this case, as Figure 7 schematically illustrates, the drift velocity $\boldsymbol{W}_d$ defined above can be interpreted as representing the motion of the pattern of dv_s/dx_s. Here the example shown in Figure 7 is a case where the particle velocity v_s has positive gradient with respect to the coordinate axis x_s. Notice that, if $W_d > v_s$, the particles behind the leading edge of the pattern dv_s/dx_s loses their momentum as the pattern passes because their velocity decreases. Here the rate of the momentum loss is $W_d \rho dv_s/dx_s$. From the viewpoint of the energy of the total system, this decrease in the momentum can be viewed as the reduction in the mechanical energy. The faster the motion of the pattern, the more energy the system loses; in the form of $W_d = \sigma_0 v_s$, $\sigma_0 > 0$, the greater the value of σ_c, the greater the energy loss. For simplicity, the x_s dependence of v_s is assumed linear in Figure 7, but the same argument holds for any x_s dependence as far as the pattern dv_s/dx_s drifts in the direction of the longitudinal force $G\boldsymbol{j}$.

The physical meaning of the dimensionless parameter σ_0 can be argued based on microscopic deformation dynamics. According to dislocation theory, plastic deformation occurs when dislocations propagate in the direction of shear stress. Here the driving force of the propagation is the local shear force, and these mobile dislocations are subject to intensive frictional force exerted by the surrounding atoms [Suzuki et al. 1991]. Based on the present field theory, this process can be explained with (28) and (40). In this context, the density on the left-hand side of these equations represents the mass of the unit volume experiencing shear strain. The first term on the right-hand side, $\nabla \times \boldsymbol{\omega}$, represents the shear force that drives mobile dislocations. The second term represents the longitudinal force exerted on the unit volume. When the material responds to the shear force elastically, as Figure 8, left, illustrates schematically, this longitudinal force is elastic force exerted by the volume behind and in front of the unit volume along the line of shear. In this situation, the longitudinal force term $G\boldsymbol{j}$ on the right-hand side of (28) represents this elastic force. When the local material yields to the shear force and starts to be deformed plastically, defects (dislocations) are generated behind or in front of the unit volume and they propagate as the shear force $\nabla \times \boldsymbol{\omega}$ is still effective. Figure 8, right, illustrates the situation schematically. In this case, the longitudinal force term $G\boldsymbol{j}$ takes the form of (40). As the defects are generated, the unit volume starts to drift and the dimensionless parameter σ_0 indicates how easily it drifts. The momentum loss discussed in Figure 7 can be interpreted as representing the energy-dissipative dynamics associated with the frictional force that the dislocations experience as they propagate. The rate of propagation of dislocations is a unique quantity of a given solid. Thus, it is natural to assume that the dimensionless parameter σ_0 is a material constant. It should be noted that the damping coefficient σ_c depends on $\nabla \cdot \boldsymbol{v}$ according to (41) and varies as the deformation status changes. As will be discussed later, an increase in $\nabla \cdot \boldsymbol{v}$ represents strain concentration and that increases the degree of energy dissipation.

Optical interferometric experiments performed on tensile-loaded metal-plate specimens [Yoshida et al. 1996; 1998] indicate that from time to time the interferometric fringe pattern shows a band structure as shown at the top of Figure 9. In this
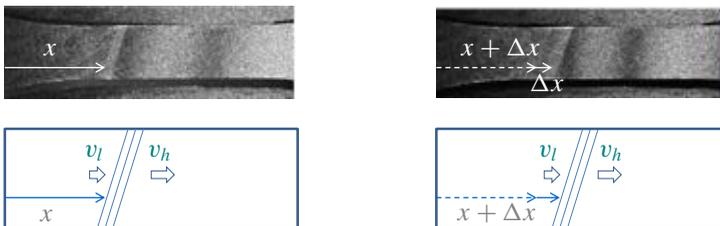

Figure 9. Developed, one-dimensional charge.

type of optical interferometry, an interferometric image of the specimen is taken continuously with a constant interval. Here the interferometric image is formed by illuminating the specimen with a pair of laser beams so that they interfere with each other on the specimen surface in such a way that the relative optical phase difference is proportional to the displacement of interest. In the case of Figure 9, the interferometer is arranged to be sensitive to in-plane displacement parallel to the tensile axis. The image taken at each time step is subtracted from the image taken at a certain time step later. The result of the subtraction yields a fringe pattern that displays a whole-field map of the differential displacement that occurs between the two time steps. Since this differential displacement is proportional to the average velocity of the duration between the two time steps, this type of fringe pattern can be viewed as representing the velocity field. Thus, hereafter, the differential displacement field derived from the fringe pattern is referred to as the velocity field. Figure 9 is a sample fringe pattern. The dark stripes seen in Figure 9 are called the interferometric fringes that represent the contours of the velocity field. Each contour indicates that the velocity along the dark fringe corresponds to a relative phase change of an integer multiple of 2π. Often a band-structured concentrated fringe pattern appears and drifts along the length of the specimen as the bottom illustrations of Figure 9 indicate schematically. In specimens free of initial stress concentration, the banded structure typically starts to appear near the yield point. In specimens with initial stress concentration, the band structure can appear at any point before the specimen yields. The higher the degree of stress concentration, the earlier it appears. It is apparent that this band structure represents plastic deformation.

Based on the observation that it represents plastic deformation, this banded structure can be interpreted as a special case of the charge $\nabla \cdot \boldsymbol{v} = dv_s/dx_s$ discussed above where the velocity field depends only on the x_s axis for the entire width of the specimen at the band-structured region. Figure 10 illustrates the situation schematically where a tensile load applied to a rectangular specimen generates a field of velocity vectors to the right and forms a pattern of $dv_s/dx_s > 0$. The three parallel lines represent the pattern of dv_s/dx_s where each line is a contour of constant velocity. The x_s axis is perpendicular to the contours, and in accordance with the above argument, as the pattern drifts in the direction of x_s, the mechanical energy of the system is dissipated. The x_p axis is set parallel to the velocity contour. As the velocity field does not depend on the x_p axis ($dv_p/dx_p = 0$), we call this
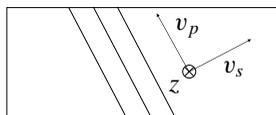

Figure 10. Developed, one-dimensional charge.

type of pattern the one-dimensional charge. As the contours cross over the width of the specimen, we call it a developed charge. Thus, the pattern shown in Figure 10 is classified as a developed, one-dimensional charge.

A developed, one-dimensional charge is often observed in tensile experiments on a low-carbon steel at the beginning of the plastic regime where the stress-strain curve shows a plateau (the yield plateau). From the temporal behavior and other features, the one-dimensional charge observed in the yield plateau has been identified as representing the phenomenon known as the Lüders band [1860]. In aluminum-alloy specimens, similar band-like interferometric fringe patterns that can also be interpreted as a one-dimensional charge are often observed in a late stage of plastic deformation. Previous studies [Yoshida et al. 1996; 1998; Yoshida and Toyooka 2001] indicate that this type of pattern represent the shear band known as the Portevin–Le Chatelier (PLC) band. A number of optical interferometric experiments indicate that, if a one-dimensional charge appears in an early or a middle stage of plastic deformation, it moves around the specimen continuously; if it appears in a late stage of plastic deformation, it tends to appear intermittently and converge its movement to a certain place of the specimen where the specimen eventually fails. We call the former type the Lüders-band-like charge and the latter the PLC-band-like charge.

From experimental observation of a developed, one-dimensional charge, we can estimate the actual value of σ_0 as follows. A previous series of tensile experiments (personal communication with T. Sasaki, 2014) on an aluminum alloy indicate that the drift velocity of the Lüders-band-like charge is proportional to the cross-head speed. In this series of experiments, the cross-head speed was set at a constant rate for each test in a range of 0.1 mm/min to 3.0 mm/min. Since only one Lüders-band-like charge appeared at a time and the number of fringes inside the charge was much greater than on the outside, we can say that the particle velocity inside the charge is approximately equal to the cross-head speed. Thus, we can approximate the magnitude of $\boldsymbol{v}$ appearing on the right-hand side of (39) ($W_d = \sigma_0 v$) by the cross-head speed; in other words, σ_0 is approximately equal to the constant of proportionality between the observed velocity of the Lüders-band-like charge (W_d) and the cross-head speed ($\approx v$). Based on this argument, the dimensionless parameter σ_0 can be estimated as $\sigma_0 \approx 3000$.

Fracture regime. Fracture can be viewed as the final stage of plastic deformation. The transition from a late stage of plastic deformation to fracture can be conveniently discussed based on the longitudinal force $G\boldsymbol{j}$. With the use of the dimensionless parameter σ_0, the plastic longitudinal force expression (38) can be put in the form

$$G\boldsymbol{j} = W_d \rho \nabla \cdot \boldsymbol{v} = \sigma_0 \rho (\nabla \cdot \boldsymbol{v}) \boldsymbol{v}. \tag{45}$$

A number of experiments indicate that, toward the end of plastic deformation, the one-dimensional charge decelerates and eventually becomes stationary at the location where the specimen fails [Yoshida et al. 1996; 1998]. The deceleration of the charge, or the corresponding decrease of the drift velocity W_d, can be understood as follows. Toward the end of the plastic regime, the material loses its capability of increasing the number of defects. In other words, the atomic arrangement does not have further room to create new dislocations. This situation eventually develops to the point where the value of σ_0 becomes zero.[3] If the external force is still active, the material needs to dissipate the energy so that the total energy is conserved. In terms of the energy-dissipative force (45), $Gj \neq 0$ and $W_d = 0$. The only possibility to make this situation to happen is $\nabla \cdot v \to \infty$. This condition can be interpreted as representing that particles flow out of the unit volume at an infinitely high rate. Obviously such a condition causes the unit volume to be empty, or the material becomes discontinuous at this location. From the viewpoint of the gauge field, the system loses the charge of symmetry that connects (the charge of the connecting field) the material to be a continuum. This is the field-theoretical definition of fracture. The above-mentioned experimental observation that a one-dimensional charge stops moving in a late stage of tensile experiment where the specimen eventually fails can be interpreted as experimental evidence of this idea about the fracture.

3. Wave dynamics of deformation and supporting experiments

3.1. *Elastic compression wave.* Substitution of (37) into (30) yields

$$\rho \frac{\partial^2 (\nabla \cdot \xi)}{\partial t^2} = \nabla \cdot \nabla (\lambda + 2G)(\nabla \cdot \xi) = (\lambda + 2G)\nabla^2 (\nabla \cdot \xi)). \qquad (46)$$

This is the equation of an elastic compression wave traveling with phase velocity

$$c_{\text{comp}} = \sqrt{\frac{\lambda + 2G}{\rho}}. \qquad (47)$$

Next, replace v with ξ, substitute (37) for Gj in (28), and take the curl of the resultant equation:

$$\rho \frac{\partial^2 (\nabla \times \xi)}{\partial t^2} = -G\nabla \times \nabla \times \omega + \nabla \times \nabla (\lambda + 2G)(\nabla \cdot \xi). \qquad (48)$$

With the mathematical identities $\nabla \times \nabla f = 0$ where $f = (\lambda + 2G)(\nabla \cdot \xi)$ and

---

[3]As discussed above, σ_0 is a material constant. The fact that it becomes zero means that at this point the material is no longer the same as before.

$\nabla \times \nabla \times \boldsymbol{\omega} = \nabla(\nabla \cdot \boldsymbol{\omega}) - \nabla^2 \boldsymbol{\omega}$, this leads to

$$\rho \frac{\partial^2 \boldsymbol{\omega}}{\partial t^2} = G \nabla^2 \boldsymbol{\omega}. \tag{49}$$

Here $\nabla \cdot \boldsymbol{\omega} = 0$ (see (26)) is used to find $\nabla(\nabla \cdot \boldsymbol{\omega}) = 0$. Equation (49) is the equation of an elastic rotational wave traveling with the phase velocity c_{shear} (see (18)). These arguments show that the field equations (23)–(26) reduce to the elastic wave equations discussed in continuum mechanics. It is interesting to note that, when the force is proportional to the differential displacement in the form of $\nabla(\nabla \cdot \boldsymbol{\xi})$ (see (37)), the longitudinal term $G \boldsymbol{j}$ vanishes when we consider the rotational effect by taking the curl of the equation because of the mathematical identity $\nabla \times \nabla f = 0$. This reflects the fact that the elastic force law is essentially longitudinal or orientation-preserving; longitudinal force does not contribute to rotational dynamics under the elastic force law.

3.2. *Plastic transverse decaying wave.* Elimination of $\boldsymbol{\omega}$ from (40) with the use of field equation (25) leads to the following wave equation that governs the velocity field $\boldsymbol{v}$:

$$\rho \frac{\partial^2 \boldsymbol{v}}{\partial t^2} - G \nabla^2 \boldsymbol{v} + \sigma_c \frac{\partial \boldsymbol{v}}{\partial t} = -G \nabla(\nabla \cdot \boldsymbol{v}). \tag{50}$$

In the pure plastic regime where the longitudinal force $G \boldsymbol{j}$ is completely energy-dissipative (see (38) and (39)) and therefore the longitudinal restoring force mechanism is absent, (50) yields decaying transverse wave solutions. The right-hand side of (50) indicates that the spatial distribution of the deformation charge is the source term of this differential equation. When the charge is uniformly distributed over the entire specimen, $\nabla(\nabla \cdot \boldsymbol{v}) = 0$ and this differential equation becomes homogeneous. In this case, we can solve it analytically and express the solution in the form

$$\boldsymbol{v}(t, \boldsymbol{r}) = \boldsymbol{v}_0 e^{-(\sigma_c/2\rho)t} \cos\left(\sqrt{\frac{G}{\rho} k^2 - \frac{\sigma_c^2}{4\rho^2}} \, t - \boldsymbol{k} \cdot \boldsymbol{r}\right). \tag{51}$$

Here $\boldsymbol{k}$ is the propagation vector.

Consider the physical meaning of the condition $\nabla(\nabla \cdot \boldsymbol{v}) = 0$. In an xyz coordinate system, the condition can be put as

$$\frac{\partial}{\partial x}\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right) = \frac{\partial^2 v_x}{\partial x^2} + \frac{\partial^2 v_y}{\partial x \partial y} + \frac{\partial^2 v_z}{\partial z \partial x} = 0, \tag{52}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right) = \frac{\partial^2 v_x}{\partial x \partial y} + \frac{\partial^2 v_y}{\partial y^2} + \frac{\partial^2 v_z}{\partial y \partial z} = 0, \tag{53}$$

$$\frac{\partial}{\partial z}\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right) = \frac{\partial^2 v_x}{\partial z \partial x} + \frac{\partial^2 v_y}{\partial y \partial z} + \frac{\partial^2 v_z}{\partial z^2} = 0. \tag{54}$$

Equations (52)–(54) indicate that under this condition the velocity field is transverse-wave-like. Here is the explanation. As a sufficient condition for (52)–(54), we can set each of the nine terms appearing after the first equals sign of the three equations to zero. Under such a condition, the first column of this part indicates that the secondary derivative of v_x is zero if the differentiation involves $\partial/\partial x$. This leads to the following generalized statements: the secondary derivative of v_i is zero if the differentiation involves $\partial/\partial x_i$ and the only surviving nonzero secondary derivative takes the form of $\partial^2 v_i/\partial x_j^2$, $i \neq j$. Note that these surviving secondary derivative terms come from the $\nabla^2 \boldsymbol{v}$ term of the wave equation (50) and not from the source term. The transverse oscillatory dynamics is due to the shear mechanism represented by $\nabla \times \boldsymbol{\omega}$. This argument indicates that the velocity wave in the pure plastic regime is essentially transverse. From this standpoint, we can put the solution (51) in the following form to express the transverse wave characteristic in a two-dimensional case. This will be compared to an experimental observation shortly below. Here k_y is the y component of $\boldsymbol{k}$, the interferometer has sensitivity in y, and the constant phase is omitted:

$$v_x(t, \boldsymbol{r}) = v_{x0} e^{-(\sigma_c/2\rho)t} \cos\left(\sqrt{\frac{G}{\rho}k^2 - \frac{\sigma_c^2}{4\rho^2}}\, t - k_y y\right). \tag{55}$$

The wave solution (55) indicates that, if σ_c is constant, the velocity field decays exponentially with the decay constant

$$\tau_c = \frac{2\rho}{\sigma_c} = \frac{2}{\sigma_0(\nabla \cdot \boldsymbol{v})}. \tag{56}$$

Based on the above argument that the dimensionless parameter σ_0 is a material-dependent constant until the solid material fractures, we can estimate the charge density $(\nabla \cdot \boldsymbol{v})$ from (56) if σ_0 and the decay constant τ_c are known. As mentioned earlier in this paper, the damping coefficient σ_c is proportional to the charge density (see (41)), and the higher the charge density, the more energy-dissipative the solid material is. Apparently, the one-dimensional, developed charge is a concentrated charge. In the context of tensile-loading, the differential velocity between the dynamic and static grips of the tensile machine is concentrated into the banded region that represents a developed charge. This situation is contrastive to an early stage of plastic deformation where the mechanical energy provided by the tensile machine is obviously dissipated via irreversible deformation but a developed charge is not present. It is interesting to estimate the charge density for both cases and compare them.

Figure 11 shows the decay characteristics of the velocity (differential displacement) field under a tensile experiment on an aluminum-alloy, thin-plate specimen. The tensile load was applied at a constant cross-head speed of 5.8×10^{-6} m/s, and
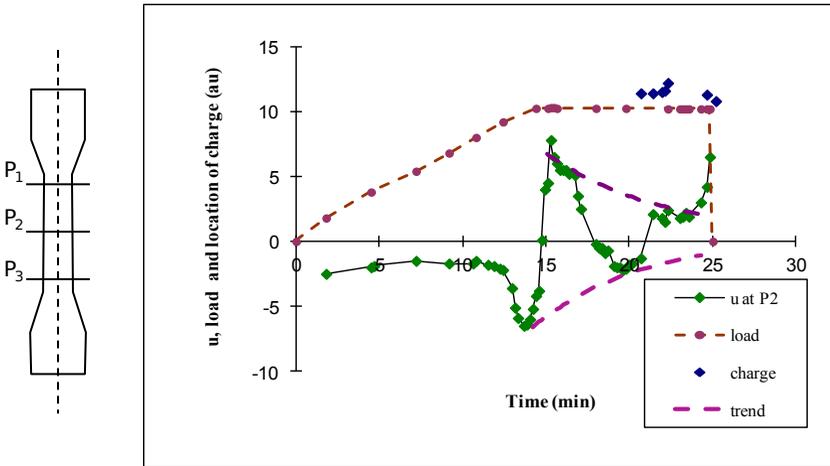
Figure 11. Decaying oscillation observed in a transverse plastic deformation wave.

the material was the same type as the above case where the dimensionless parameter σ_0 was estimated to be 3000 based on the proportionality between the drift velocity of the developed charge and the cross-head speed. The graph in Figure 11 plots the velocity vector component perpendicular to the tensile axis (the transverse velocity) evaluated at the reference point P_2 shown in the left margin of the figure. Also plotted are the load recorded by the tensile machine and the location of PLC-like charge that started to appear toward the end of the plastic regime. The horizontal axis is the time elapsed from the beginning of the tensile-loading. The oscillatory behavior of the transverse velocity starts near the yield point. In fact, observation of the transverse velocity at the other reference points (P_1 and P_3) indicates that the oscillatory behavior propagates as a transverse wave [Yoshida et al. 1999]. Thus, it is reasonable to interpret the plot in Figure 11 as the decay characteristic of the transverse velocity wave in the plastic regime. Since the developed charge does not appear until the oscillatory behavior fades out, it is expected that the charge density is lower than that typically observed under the existence of a PLC-like charge.

From the trend of the oscillation peaks seen in Figure 11, the decay time constant of this case can be estimated as $6.7 \, \text{min} = 400 \, \text{s}$. Substituting this value into the left-hand side of (56) and using $\sigma_0 = 3000$ for the aluminum-alloy case, we can estimate the charge density during this decay process as $(\nabla \cdot \boldsymbol{v}) = 2/(400 \times 3000) = 1.7 \times 10^{-6} \, 1/\text{s}$. On the other hand, the charge density when the PLC-like charge appears can be estimated as follows. Using the same logic as above that the velocity of the leading edge (the edge closer to the dynamic grip) of the charge is approximately equal to that of the dynamic grip and the velocity of the tailing edge is zero, and using the band width of $5.2 \, \text{mm}$ along the tensile axis, we

can evaluate the charge density $dv_y/dy = 5.8 \times 10^{-6}\,\text{m/s}\,/5.2 \times 10^{-3}\,\text{mm} = 1.1 \times 10^{-3}\,\text{1/s}$. As expected, this charge density is three orders of magnitude higher than $(\nabla \cdot \boldsymbol{v}) = 2/(400 \times 3000) = 1.7 \times 10^{-6}\,\text{1/s}$ evaluated at the beginning of the oscillatory behavior of the transverse velocity.

3.3. Solitary wave in plastic regime. In the transitional stage from the elastic regime to the plastic regime, optical interferometric experiments often show a band-structured, interferometric fringe pattern that can be interpreted as the one-dimensional charge expressed by (42) and illustrated by Figures 9 and 10. From the similarity in various behaviors, as mentioned above, this type of one-dimensional charge can be interpreted as representing the same phenomenon as the Lüders band [Yoshida et al. 2005]. Among these behaviors, the following two are interesting from the viewpoint of dynamics: (a) their drift velocity is proportional to the tensile rate and (b) the stress remains practically the same while they drift. Mertens et al. [1997] have made detailed analyses on dynamic behaviors of the Lüders band. They explain the mechanism of the phenomenon as the propagation of mobile dislocations at the plastic deformation front weakens the neighboring areas and the resultant deformation creates new dislocations. They also explain that the deformation at the front creates a strain jump that is roughly constant during the drift of the band and therefore its drift velocity is proportional to the pulling rate. Here an attempt is made to explain the behaviors (a) and (b) based on the present field theory.

Figure 10 illustrates schematically that one can characterize the velocity contours inside a one-deformation charge of this type as

$$\frac{\partial}{\partial x_s} \neq 0, \tag{57}$$

$$\frac{\partial}{\partial x_p} = 0. \tag{58}$$

In this two-dimensional picture, the one-dimensional, developed charge is expressed as

$$\nabla \cdot \boldsymbol{v} = \frac{\partial v_p}{\partial x_p} + \frac{\partial v_s}{\partial x_s} = \frac{dv_s}{dx_s}. \tag{59}$$

Similarly, the volume expansion appearing in the elastic longitudinal force expression (37) is

$$\nabla \cdot \boldsymbol{\xi} = \frac{\partial \xi_p}{\partial x_p} + \frac{\partial \xi_s}{\partial x_s} = \frac{d\xi_s}{dx_s}. \tag{60}$$

The rotation vector has only the z component:

$$\boldsymbol{\omega} = \omega_z \hat{z} = \left(\frac{\partial \xi_s}{\partial x_p} - \frac{\partial \xi_p}{\partial x_s} \right) \hat{z}. \tag{61}$$

Thereby, its rotation can be expressed as

$$\nabla \times \boldsymbol{\omega} = \frac{\partial \omega_z}{\partial x_s}\hat{x}_p - \frac{\partial \omega_z}{\partial x_p}\hat{x}_s = \frac{\partial \omega_z}{\partial x_s}\hat{x}_p, \tag{62}$$

where the z axis is perpendicular to the x_s-x_p plane and condition (58) is used in going through the last equal sign.

Equations (59), (60) and (62) respectively indicate that inside the one-dimensional charge along the x_s axis both the longitudinal plastic force proportional to $\nabla \cdot \boldsymbol{v}$ and elastic force proportional to $\nabla \cdot \boldsymbol{\xi}$ are potentially nonzero, and the shear force component $(G\nabla \times \boldsymbol{\omega})_s$ is zero. From the fact that the entire specimen is being pulled by the tensile machine, it is apparent that this elastic longitudinal force is tensile. The fact that the stress recorded by the tensile machine does not increase indicates that, as this tensile force stretches the banded region, stress drops occur presumably associated with the creation of new mobile dislocations. Thus, it is reasonable to assume that the elastic behavior is confined within the banded region. This argument leads to the following physical model. An elastic medium isolated to the location of the band moves in the entire specimen due to the longitudinal elastic force. The elastic dynamics is not transferred to outside of the banded region due to the plastic deformation associated with the creation of dislocations at the front. As the charge represented by the band region moves, the plastic longitudinal force causes energy dissipation. The coexistence of the elastic and plastic deformation makes total sense as this phenomenon takes place in the transitional stage from the elastic regime to the plastic regime.

Based on the above explanation, we can start a quantitative argument from the equation of motion of the elastic block (called the block) confined in the banded region. The net elastic force acting on the block is the differential force between the front and back surfaces of the block. At each surface, the elastic force is proportional to the local stretch, as Figure 12 illustrates schematically:

$$\eta(x_s) = \frac{\partial \xi_s(x_s)}{\partial x_s}\delta x_s. \tag{63}$$

Here $\eta(x_s)$ is the stretch at x_s, $\xi(x_s)$ is the displacement at the same point, and δx_s is the infinitesimal width of the plane at x_s. The displacement of the block from its equilibrium position X is the differential displacement of its front and back ends:

$$X = \frac{\partial \eta}{\partial x_s}\Delta x_s = \frac{\partial}{\partial x_s}\left(\frac{\partial \xi_s}{\partial x_s}\delta x_s\right)\Delta x_s = \frac{\partial^2 \xi_s}{\partial x_s^2}(\delta x_s \Delta x_s), \tag{64}$$
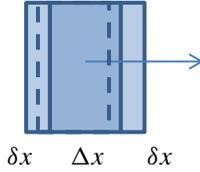
where Δx_s is the width of the block. The corresponding elastic energy is

$$U = \tfrac{1}{2}kX^2 = \tfrac{1}{2}k\left(\frac{\partial^2 \xi_s}{\partial x_s^2}\right)^2(\delta x_s \Delta x_s)^2 = \frac{SE}{2}\left(\frac{\partial^2 \xi_s}{\partial x_s^2}\right)^2 \delta x_s(\Delta x_s)^2, \tag{65}$$

$\eta(x_s)$: displacement from equilibrium due to elasticity

$\eta(x) = (\partial \xi / \partial x)\big|_x \delta x$ $\qquad\qquad$ $\eta(x + \Delta x) = (\partial \xi / \partial x)\big|_{x+\Delta x} \delta x$

X: band's displacement from equilibrium

$\delta x \quad \Delta x \quad \delta x$ $\qquad\qquad$ $X = \dfrac{\partial \eta}{\partial x} \Delta x$

$\longrightarrow x_s \equiv x$

Figure 12. Elastic force acting on the isolated region that can be identified as a one-dimensional charge.

where E is the elastic modulus associated with the longitudinal dynamics (corresponding to the stiffness or the Young's modulus in the elastic regime),[4] S is the cross-sectional area and $k\delta x_s = SE$. This leads to the Lagrangian density of

$$L_{\text{charge}} = \frac{U}{S\Delta x_s} = \frac{E}{2}\left(\frac{\partial^2 \xi_s}{\partial x_s^2}\right)^2 (\delta x_s \Delta x_s) = \frac{E}{2}(\partial_{x_s}^2 \xi_s)^2 (\delta x_s \Delta x_s). \qquad (66)$$

Thus, the corresponding term of the Euler–Lagrangian equation of motion is

$$\partial_{x_s}^2\left(\frac{\partial L_{\text{charge}}}{\partial(\partial_{x_s}^2 \xi_s)}\right) = E\partial_{x_s}^2(\partial_{x_s}^2 \xi_s)(\delta x_s \Delta x_s) = E\partial_{x_s}^4 \xi_s(\delta x_s \Delta x_s). \qquad (67)$$

Writing the traveling band in the form $\xi(x_s, t) = \xi(x_s - c_w t)$, we can replace one of the spatial derivatives with a temporal derivative by $\partial_{x_s} = -\partial_t \xi_s / c_w = -v_s / c_w$ (see (10)). With this, (67) becomes

$$\partial_{x_s}^2\left(\frac{\partial L_{\text{charge}}}{\partial \partial_{x_s}^2 \xi}\right) = -\frac{E}{c_w}\partial_{x_s}^3(\partial_t \xi_s)(\delta x_s \Delta x_s). \qquad (68)$$

When a developed charge appears, the material loses the elastic restoring force associated with $(\nabla \times \boldsymbol{\omega})_s$, and the longitudinal energy-dissipative force $\sigma_0 \rho (\nabla \cdot \boldsymbol{v})v = \sigma_0 \rho \partial v_s / \partial x_s v_s$ is active at the front where plastic deformation causes the stress drop. So (40) can be written in the form

$$\rho \frac{\partial v_s}{\partial t} = -\sigma_0 v_s \rho \frac{\partial v_s}{\partial x_s} - \frac{E\delta x_s \Delta x_s}{c_w}\partial_{x_s}^3(\partial_t \xi_s), \qquad (69)$$

---

[4] E is the stiffness associated with the normal stress or the longitudinal effect. In the plastic regime, or when the defect density is substantial, this value becomes lower than in the elastic regime. In the present context, it should be differentiated from the Young's modulus of the elastic regime.

which can further be rewritten in a form more familiar as the Korteweg–de Vries equation [Maugin 2011]:

$$\partial_t v + \sigma_0 v \partial_x v + \frac{E \delta x_s \Delta x_s}{\rho c_w} \partial_x^3 v = 0. \tag{70}$$

Here, for clarity, the subscript s has been omitted from the variables. As is well known, (70) yields the form of solution

$$v(x, t) = a \operatorname{sech}^2(b(x - c_w t)), \tag{71}$$

where

$$c_w = \frac{\sigma_0 a}{3}, \tag{72}$$

$$b^2 = \left(\frac{\sigma_0 a}{3}\right)^2 \frac{\rho}{4 E \delta x_s \Delta x_s}. \tag{73}$$

In condition (72), a is the amplitude of the velocity wave $v(x, t)$. It is reasonable to consider that this amplitude is proportional to the pulling rate of the specimen. This explains why the one-dimensional charge, and hence the Lüders band, drifts at a velocity proportional to the pulling rate. Condition (73) indicates that the width of the banded region is proportional to the square root of the elastic modulus E. This indicates that, as the material loses its elasticity with the development of plastic deformation, the one-dimensional charge tends to be narrower. This is consistent with experimental observations [Yoshida et al. 1996; 1998] and can be interpreted as the degree of stress concentration increasing with the development of plastic deformation.

4. Conclusions

Based on the field-theoretical approach associated with the local symmetry of linear elastic law, the dynamics of deformation and fracture has been discussed. Various conventionally known phenomena of deformation have been explained from the field-theoretical viewpoint. The concept of deformation charge has been introduced based on the analogy to electrodynamics and used to explain the energy-dissipative nature of plastic deformation. Decaying transverse wave characteristics of plastic deformation and solitary wave characteristics of the transitional stage from the elastic regime to the plastic regime have been discussed and compared to experimental results.

References

[Aitchison and Hey 1989] I. J. R. Aitchison and A. J. G. Hey, *Gauge theories in particle physics*, 2nd ed., Institute of Physics, Bristol, 1989.

[Auffray et al. 2015] N. Auffray, F. dell'Isola, V. A. Eremeyev, A. Madeo, and G. Rosi, "Analytical continuum mechanics à la Hamilton–Piola least action principle for second gradient continua and capillary fluids", *Math. Mech. Solids* **20**:4 (2015), 375–417.

[Fleck and Hutchinson 1997] N. A. Fleck and J. W. Hutchinson, "Strain gradient plasticity", *Adv. Appl. Mech.* **33** (1997), 295–361.

[Griffith 1921] A. A. Griffith, "The phenomena of rupture and flow in solids", *Philos. T. Roy. Soc. A* **221**:587 (1921), 163–198.

[Hill 1998] R. Hill, *The mathematical theory of plasticity*, Oxford University, 1998.

[Ichinose et al. 2006] K. Ichinose, S. Yoshida, K. Gomi, K. Taniuchi, K. Fukuda, and H. Ishii, "Detection of crack initiation by observations of free surface condition", *J. ASTM Int.* **3**:6 (2006), JAI13189.

[Irwin 1948] G. R. Irwin, "Fracture dynamics", pp. 147–166 in *Fracturing of metals*, edited by F. Jonassen et al., American Society for Metals, Cleveland, 1948.

[Javili et al. 2013] A. Javili, F. dell'Isola, and P. Steinmann, "Geometrically nonlinear higher-gradient elasticity with energetic boundaries", *J. Mech. Phys. Solids* **61**:12 (2013), 2381–2401.

[Landau and Lifshitz 1986] L. D. Landau and E. M. Lifshitz, *Course of theoretical physics, VII: Theory of elasticity*, 3rd ed., Butterworth–Heinemann, Oxford, 1986.

[Lubliner 2008] J. Lubliner, *Plasticity theory*, Dover, Mineola, NY, 2008.

[Lüders 1860] W. Lüders, "Ueber die Aeußerung der Elasticität an stahlartigen Eisenstäben und Stahlstäben, und über eine beim Biegen solcher Stäbe beobachtete Molecularbewegung", *Polytech. J. Dingler* **155**:8 (1860), 18–22.

[Madeo et al. 2013] A. Madeo, F. dell'Isola, and F. Darve, "A continuum model for deformable, second gradient porous media partially saturated with compressible fluids", *J. Mech. Phys. Solids* **61**:11 (2013), 2196–2211.

[Maugin 2011] G. A. Maugin, "Solitons in elastic solids (1938–2010)", *Mech. Res. Commun.* **38**:5 (2011), 341–349.

[Mertens et al. 1997] F. Mertens, S. V. Franklin, and M. Marder, "Dynamics of plastic deformation fronts in an aluminum alloy", *Phys. Rev. Lett.* **78**:23 (1997), 4502–4505.

[Mindlin 1965] R. D. Mindlin, "Second gradient of strain and surface-tension in linear elasticity", *Int. J. Solids Struct.* **1**:4 (1965), 417–438.

[Spencer 1980] A. J. M. Spencer, *Continuum mechanics*, Longman, London, 1980.

[Suzuki et al. 1991] T. Suzuki, S. Takeuchi, and H. Yoshinaga, *Dislocation dynamics and plasticity*, Series in Materials Science **12**, Springer, Berlin, 1991.

[Toupin 1964] R. A. Toupin, "Theories of elasticity with couple-stress", *Arch. Rational Mech. Anal.* **17**:2 (1964), 85–112.

[Yoshida 2000] S. Yoshida, "Consideration on fracture of solid-state materials", *Phys. Lett. A* **270**:6 (2000), 320–325.

[Yoshida 2008] S. Yoshida, "Dynamics of plastic deformation based on restoring and energy dissipative mechanisms in plasticity", *Phys. Mesomech.* **11**:3–4 (2008), 137–143.

[Yoshida 2011] S. Yoshida, "Scale-independent approach to deformation and fracture of solid-state materials", *J. Strain Anal. Eng.* **46**:5 (2011), 380–388.

[Yoshida 2015] S. Yoshida, *Deformation and fracture of solid-state materials: field theoretical approach and engineering applications*, Springer, New York, 2015.

[Yoshida and Toyooka 2001] S. Yoshida and S. Toyooka, "Field theoretical interpretation on dynamics of plastic deformation: Portevin–Le Chatelie effect and propagation of shear band", *J. Phys. Condens. Mat.* **13**:31 (2001), 6741–6757.

[Yoshida et al. 1996] S. Yoshida, Suprapedi, R. Widiastuti, M. Pardede, S. Hutagalong, J. S. Marpaung, A. F. Muhardy, and A. Kusnowo, "Direct observation of developed plastic deformation and its application to nondestructive testing", *Jpn. J. Appl. Phys.* **35**:7A (1996), L854–L857.

[Yoshida et al. 1998] S. Yoshida, Muchiar, I. Muhamad, R. Widiastuti, and A. Kusnowo, "Optical interferometric technique for deformation analysis", *Opt. Express* **2**:13 (1998), 516–530.

[Yoshida et al. 1999] S. Yoshida, B. Siahaan, M. H. Pardede, N. Sijabat, H. Simangunsong, T. Simbolon, and A. Kusnowo, "Observation of plastic deformation wave in a tensile-loaded aluminum-alloy", *Phys. Lett. A* **251**:1 (1999), 54–60.

[Yoshida et al. 2005] S. Yoshida, H. Ishii, K. Ichinose, K. Gomi, and K. Taniuchi, "An optical interferometric band as an indicator of plastic deformation front", *J. Appl. Mech.* **72**:5 (2005), 792–794.

SANICHIRO YOSHIDA: syoshida@selu.edu
Department of Chemistry and Physics, Southeastern Louisiana University, Hammond, LA 70402, United States
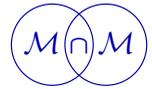
ON THE CONSTITUTIVE EQUATIONS
OF VISCOELASTIC MICROPOLAR PLATES AND SHELLS
OF DIFFERENTIAL TYPE

HOLM ALTENBACH AND VICTOR A. EREMEYEV

Within the framework of the micropolar theory of continuum we discuss the constitutive equations of viscoelastic micropolar thin-walled structures, i.e. viscoelastic micropolar plates and shells. Starting from the linear viscoelastic micropolar continuum and using the correspondence principle of the linear viscoelasticity we extend the procedure of reduction of three-dimensional equilibrium equations of elastic shell-like solids to the case of viscoelastic behavior. We restricted ourselves by constitutive equations of differential type. In other words, we consider both 2D and 3D constitutive equations which are linear dependencies between certain set of time derivatives of stress and strain measures.

1. Introduction

The model of the Cosserat or micropolar continuum has recently found new applications in the modeling of the behavior of materials and structures with complex inner structure; see [Eremeyev et al. 2013] and references therein. Since the seminal paper [Ericksen and Truesdell 1958] the Cosserat model has found numerous applications in construction of various generalized models for beams, plates, and shells; see the review and bibliography in [Altenbach et al. 2009]. Within the framework of the direct approach of Ericksen and Truesdell, the shell is modeled as a deformable surface at each point of which a set of directors additionally is attached. In the literature are also known theories of plates and shells based on the reduction of three-dimensional micropolar continuum equations; see for example [Eringen 1999; Reissner 1977; Sargsyan 2011; Zubov 2009; Steinberg and Kvasov 2013; Altenbach et al. 2009], where various averaging procedures in the thickness direction together with the approximation of the displacements and rotations or the force and moment stresses in the thickness direction are applied. In major cases these considerations are restricted by the elastic behavior. For the linear theory of

viscoelasticity the application of the correspondence principle gives the possibility to derive the theory of viscoelasticity in the case of thin-walled structures such as plates and shells.

In this paper we extend the through-the thickness integration procedure applied in [Altenbach and Eremeyev 2009] to viscoelastic micropolar plates and shells. The interest to the theory of viscoelastic micropolar thin-walled structures is based on prospective applications of this theory to the mechanics of plates and shells made of materials with complex inner structure, for example cellular materials and foams for which the micropolar model is used; see [Diebels and Steeb 2003; Lakes 1986] among others.

We consider here the variant of plates and shell theory based on six kinematically independent variables, namely translations and rotations. This theory can be derived using a direct approach [Eremeyev et al. 2013; Eremeyev and Zubov 2008; Rubin 2000] or based on the reduction of the three-dimensional motion equations [Chróścielewski et al. 2004; Libai and Simmonds 1998]. The mathematical study of boundary-value problems was performed in [Bîrsan and Neff 2013; 2014; Eremeyev and Lebedev 2011], while various solutions and finite-element calculations are presented in [Chróścielewski et al. 2004; 2010; 2011; Eremeyev and Zubov 2008].

2. Basic relations of the viscoelastic Cosserat continuum.

Following [Eringen 1999] we recall the governing equations of the linear micropolar viscoelasticity. Let the micropolar body occupies the domain $\mathcal{V} \in \mathbb{R}^3$. The infinitesimal deformations of the micropolar media are described by two vectorial fields. The first one is the vector of translation $\boldsymbol{u}$ and the second field is the vector of microrotation $\boldsymbol{\vartheta}$ given as vector-functions of the position vector $\boldsymbol{x}$ and time t. From the physical point of view, $\boldsymbol{u}$ describes an displacement of the particle of a micropolar body while $\boldsymbol{\vartheta}$ corresponds to the particle rotation. The quasistatic deformations of a micropolar body are described by the equilibrium equations

$$\nabla \cdot \boldsymbol{\sigma} + \rho \boldsymbol{f} = \boldsymbol{0}, \quad \nabla \cdot \boldsymbol{\mu} + \boldsymbol{\sigma}_\times + \rho \boldsymbol{\ell} = \boldsymbol{0}, \qquad \boldsymbol{x} \in \mathcal{V}, \tag{1}$$

where ∇ is the three-dimensional nabla operator, $\boldsymbol{f}$ and $\boldsymbol{\ell}$ are the mass forces and the mass couples vectors, ρ is the density, $\boldsymbol{\sigma}$ and $\boldsymbol{\mu}$ are the stress tensor and the couple stress tensor, respectively, and $\boldsymbol{\sigma}_\times$ denotes the vectorial invariant of the second-order tensor $\boldsymbol{\sigma}$, see [Lebedev et al. 2010; Wilson 1901]. Equation $(1)_1$ represents the local form of the balance of momentum while Eq. $(1)_2$ is the balance of moment of momentum.

The static boundary conditions have the following form

$$\boldsymbol{n} \cdot \boldsymbol{\sigma} = \boldsymbol{t}^0, \quad \boldsymbol{n} \cdot \boldsymbol{\mu} = \boldsymbol{m}^0 \quad \text{at } S_f. \tag{2}$$

Here t^0 and m^0 are the surface forces and the surface couples acting on the part of the surface S_f of the micropolar body, $S = S_u \cup S_f \equiv \partial \mathcal{V}$. The kinematic boundary conditions consist of the following relations

$$u = u^0, \quad \vartheta = \vartheta^0 \quad \text{at } S_u, \tag{3}$$

where u^0 and ϑ^0 are given functions at S_u. Let us note that since the displacements and rotations are independent kinematic variables, Dirichlet and Neumann boundaries for the displacements and rotations are different, in general. As a result, the kinematic and static boundary conditions take the form

$$\begin{aligned} n \cdot \sigma = t^0 &\quad \text{at } S_f, \qquad u = u^0 \quad \text{at } S_u, \\ n \cdot \mu = m^0 &\quad \text{at } S_m, \qquad \vartheta = \vartheta^0 \quad \text{at } S_\theta, \end{aligned} \tag{4}$$

where $S = S_u \cup S_f = S_m \cup S_\theta$ are two decompositions of S. For simplicity in what follows we use the same boundaries that is $S_u = S_\theta$ and $S_f = S_m$. Obviously, other mixed boundary conditions are also possible [Eremeyev et al. 2013].

The linear strain measures, i.e. the linear stretch tensor ε and the linear wryness tensor $\mathit{æ}$, are given by the relations

$$\varepsilon = \nabla u + \vartheta \times \mathbf{I}, \quad \mathit{æ} = \nabla \vartheta, \tag{5}$$

where $\mathbf{I}$ is the unit three-dimensional tensor and $\times$ the cross product.

Let us consider the viscoelastic micropolar isotropic material. For the sake of simplicity we restrict ourselves by the constitutive equations of the differential type. This means that the constitutive equations are

$$P_0(\partial_t) \sigma = P_1(\partial_t) \varepsilon + P_2(\partial_t) \varepsilon^T + P_3(\partial_t) \mathbf{I} \operatorname{tr} \varepsilon, \tag{6}$$

$$Q_0(\partial_t) \mu = Q_1(\partial_t) \mathit{æ} + Q_2(\partial_t) \mathit{æ}^T + Q_3(\partial_t) \mathbf{I} \operatorname{tr} \mathit{æ}. \tag{7}$$

In (6) and (7) ∂_t stands for the derivative with respect to time; P_0, P_1, P_2, P_3, Q_0, Q_1, Q_2, Q_3 are the polynomials.

Using the Laplace transform of a function $f(t)$, given by

$$\bar{f}(p) = \int_0^\infty f(t) e^{-pt} \, dt,$$

the constitutive equations (6) and (7) become

$$\bar{\sigma} = \lambda \mathbf{I} \operatorname{tr} \bar{\varepsilon} + \mu \bar{\varepsilon}^T + (\mu + \kappa) \bar{\varepsilon}, \quad \bar{\mu} = \alpha \mathbf{I} \operatorname{tr} \bar{\mathit{æ}} + \beta \bar{\mathit{æ}}^T + \gamma \bar{\mathit{æ}}, \tag{8}$$

where

$$\lambda = \frac{P_3(p)}{P_0(p)}, \quad \mu = \frac{P_2(p)}{P_0(p)}, \quad \kappa = \frac{P_1(p) - P_2(p)}{P_0(p)}, \tag{9}$$

$$\alpha = \frac{Q_3(p)}{Q_0(p)}, \quad \beta = \frac{Q_2(p)}{Q_0(p)}, \quad \gamma = \frac{Q_1(p)}{Q_0(p)} \tag{10}$$

are the Laplace transforms of the relaxation functions of the viscoelastic micropolar material. In fact, (8) with (9), (10) coincide up to notations with the constitutive equations of an isotropic linear micropolar solid, see [Eringen 1999].

Substituting (8) into (1) one may derive the equilibrium equations in terms of the kinematical fields $\bar{\boldsymbol{u}}$ and $\bar{\boldsymbol{\vartheta}}$. For homogeneous micropolar bodies these equations are

$$\begin{aligned}
(\lambda + \mu)\nabla\nabla \cdot \bar{\boldsymbol{u}} + (\mu + \kappa)\nabla \cdot \nabla\bar{\boldsymbol{u}} + \kappa\nabla \times \bar{\boldsymbol{\vartheta}} + \rho\bar{\boldsymbol{f}} = \boldsymbol{0}, \\
(\alpha + \beta)\nabla\nabla \cdot \bar{\boldsymbol{\vartheta}} + \gamma\nabla \cdot \nabla\bar{\boldsymbol{\vartheta}} + \kappa\nabla \times \bar{\boldsymbol{u}} - 2\kappa\bar{\boldsymbol{\vartheta}} + \rho\bar{\boldsymbol{\ell}} = \boldsymbol{0}.
\end{aligned} \tag{11}$$

Remark. In a similar manner other linear viscoelastic constitutive equations (integral equation, complex moduli equations) can be introduced and applied to the representation of the viscoelastic micropolar behavior.

In what follows we consider the reduction of three-dimensional equilibrium equations to the two-dimensional ones and discuss the corresponding two-dimensional constitutive equations.

3. Micropolar plate and shell equations

Within the framework of the linear theory of micropolar plates and shells, also called the six-parameter theory of shells, we consider a micropolar shell as a two-dimensional Cosserat continuum, i.e., as a deformable surface $\mathcal{M}$ with six degrees of freedom. Each material point of the surface is kinematically similar to an infinitesimal rigid body with three translational and three rotational degrees of freedom. Hence, the deformations of the micropolar plate or shell are described by the translation vector $\boldsymbol{v}$ and the rotation vector $\boldsymbol{\theta}$ which are defined at the base surface $\mathcal{M}$. Using the direct approach the basics of the linear theory of micropolar shells are summarized in [Eremeyev and Zubov 2008; Lebedev et al. 2010]; see also Appendix D in [Eremeyev et al. 2013]. The governing equations of the micropolar shells and plates coincide with the relations of the general 6-parameter nonlinear shell theory presented in [Chróścielewski et al. 2004; Libai and Simmonds 1998] in the case of small deformations derived using the through-the-thickness integration procedure.

The balance of momentum and the balance of moment of momentum are formulated as follows

$$\nabla_s \cdot \mathbf{T} + \boldsymbol{q} = \boldsymbol{0}, \quad \nabla_s \cdot \mathbf{M} + \mathbf{T}_\times + \boldsymbol{c} = \boldsymbol{0}, \tag{12}$$

where $\boldsymbol{q}$ and $\boldsymbol{c}$ are the surface loads (forces and moments), $\mathbf{T}$ and $\mathbf{M}$ are the resultant force stress and couple stress tensors, and ∇_s is the surface nabla differential

operator. The tensors $\mathbf{T}$ and $\mathbf{M}$ have the properties

$$\boldsymbol{n} \cdot \mathbf{T} = \mathbf{0}, \quad \boldsymbol{n} \cdot \mathbf{M} = \mathbf{0}, \tag{13}$$

where $\mathbf{n}$ is the unit vector of normal to $\mathcal{M}$. Hence, $\mathbf{T}$ and $\mathbf{M}$ take the form

$$\mathbf{T} = T_{\alpha\beta} \boldsymbol{i}^\alpha \otimes \boldsymbol{i}^\beta + T_{\alpha 3} \boldsymbol{i}^\alpha \otimes \boldsymbol{n}, \quad \mathbf{M} = M_{\alpha\beta} \boldsymbol{i}^\alpha \otimes \boldsymbol{i}^\beta + M_{\alpha 3} \boldsymbol{i}^\alpha \otimes \boldsymbol{n} \quad (\alpha, \beta = 1, 2), \tag{14}$$

where $\boldsymbol{i}^\alpha$ are the base vectors on the shell surface. Obviously, in this theory of shells the action of the drilling moment $M_{\alpha 3}$ is taken into account. For example, such possibility may be useful to describe the interaction of the shell and the rigid body or for the description of the deformations of multifolded plates, see [Chróścielewski et al. 2004].

The static and kinematic boundary conditions take the form

$$\begin{aligned} \boldsymbol{v} \cdot \mathbf{T} = \boldsymbol{t}_s^*, \quad \boldsymbol{v} \cdot \mathbf{M} = \boldsymbol{m}_s^* \quad & \text{along } \mathcal{C}_f, \\ \boldsymbol{v} = \boldsymbol{v}^0, \quad \boldsymbol{\theta} = \boldsymbol{\theta}^0 \quad & \text{along } \mathcal{C}_u. \end{aligned} \tag{15}$$

Here $\boldsymbol{v}$ is the vector of the unit normal to $\mathcal{C}_f$, $\boldsymbol{v} \cdot \boldsymbol{n} = 0$, $\boldsymbol{t}_s^*$ and $\boldsymbol{m}_s^*$ are external force and couple vectors acting along the boundary $\mathcal{C}_f$, while $\boldsymbol{v}^0$ and $\boldsymbol{\theta}^0$ are given functions describing the displacements and rotations of the boundary $\mathcal{C}_u$, respectively, $\mathcal{C} = \mathcal{C}_u \cup \mathcal{C}_f \equiv \partial\mathcal{M}$. For simplicity we use here again the same boundaries for Dirichlet and Neumann conditions for displacements and rotations.

The linear surface strain measures are

$$\boldsymbol{\epsilon} = \nabla_s \boldsymbol{v} + \mathbf{A} \times \boldsymbol{\theta}, \quad \boldsymbol{\kappa} = \nabla_s \boldsymbol{\theta}, \tag{16}$$

where $\mathbf{A} \equiv \mathbf{I} - \boldsymbol{n} \otimes \boldsymbol{n}$ is the two-dimensional or surface unit tensor.

We restrict ourselves again by the constitutive equations for $\mathbf{T}$ and $\mathbf{M}$ of differential type. We assume the following relations

$$A_0(\partial_t)\mathbf{T} = \mathbf{A} A_1(\partial_t) \operatorname{tr} \boldsymbol{\epsilon}_\| + A_2(\partial_t)\boldsymbol{\epsilon}_\|^T + A_3(\partial_t)\boldsymbol{\epsilon}_\| + A_4(\partial_t)\boldsymbol{\epsilon} \cdot \boldsymbol{n} \otimes \boldsymbol{n}, \tag{17}$$

$$B_0(\partial_t)\mathbf{M} = \mathbf{A} B_1(\partial_t) \operatorname{tr} \boldsymbol{\kappa}_\| + B_2(\partial_t)\boldsymbol{\kappa}_\|^T + B_3(\partial_t)\boldsymbol{\kappa}_\| + B_4(\partial_t)\boldsymbol{\kappa} \cdot \boldsymbol{n} \otimes \boldsymbol{n}, \tag{18}$$

Here $\boldsymbol{\epsilon}_\| = \boldsymbol{\epsilon} \cdot \mathbf{A}$, $\boldsymbol{\kappa}_\| = \boldsymbol{\kappa} \cdot \mathbf{A}$, and $A_k(p)$, $B_k(p)$, $k = 0, 1, 2, 3$, are polynomials.

We call the constitutive equations for the shell (17) and (18) the Maxwell-type, if $A_0 = 1 + a_0 p$, $B_0 = 1 + b_0 p$, where $1/a_0$ and $1/b_0$ are the relaxation time for the stress resultants and for couple stresses, respectively, and if other polynomials are constants. We call (17) and (18) the Voigt-type constitutive equation, if $A_0 = B_0 = 1$ while other polynomials are linear functions of p.

Using the Laplace transform these constitutive equations become

$$\overline{\mathbf{T}} = \alpha_1 \mathbf{A} \operatorname{tr} \overline{\boldsymbol{\epsilon}}_\| + \alpha_2 \overline{\boldsymbol{\epsilon}}_\|^T + \alpha_3 \overline{\boldsymbol{\epsilon}}_\| + \alpha_4 \overline{\boldsymbol{\epsilon}} \cdot \boldsymbol{n} \otimes \boldsymbol{n}, \tag{19}$$

$$\overline{\mathbf{M}} = \beta_1 \mathbf{A} \operatorname{tr} \overline{\boldsymbol{\kappa}}_\| + \beta_2 \overline{\boldsymbol{\kappa}}_\|^T + \beta_3 \overline{\boldsymbol{\kappa}}_\| + \beta_4 \overline{\boldsymbol{\kappa}} \cdot \boldsymbol{n} \otimes \boldsymbol{n}, \tag{20}$$

where

$$\alpha_i = \frac{A_i(p)}{A_0(p)}, \quad \beta_i = \frac{B_i(p)}{B_0(p)}, \qquad i = 1, 2, 3, 4.$$

The final step in the theory of micropolar viscoelastic shells is the construction of the polynomials A_k and B_k on the base of the dimension reduction of the three-dimensional continuum or by experiments. Further we consider the derivation of the relaxation functions α_i and β_i using the through-the-thickness integration procedure presented in [Altenbach and Eremeyev 2009; 2010].

4. Reduction of constitutive equations from three to two dimensions

For simplicity let us consider the undeformed plane geometry; i.e., we restrict ourselves to the theory of plates. The case of shells can be discussed by similar way. Using the correspondence principle [Christensen 1971] from [Altenbach and Eremeyev 2009] it follows

$$\overline{\mathbf{T}} = \langle \mathbf{A} \cdot \overline{\boldsymbol{\sigma}} \rangle, \quad \overline{\mathbf{M}} = \langle \mathbf{A} \cdot \overline{\boldsymbol{\mu}} \rangle - \langle \mathbf{A} \cdot z \overline{\boldsymbol{\sigma}} \times \boldsymbol{n} \rangle, \tag{21}$$

where z is the coordinate along the thickness coordinate axis. The averaging operator $\langle \cdot \rangle$ is defined as follows

$$\langle f \rangle = \int_{-h/2}^{h/2} f(z) \, \mathrm{d}z.$$

For derivation of (21) we refer to the Appendix.

As a result the two-dimensional material parameters are given by the relations

$$\overline{\alpha}_1 = \Lambda h \equiv \frac{\overline{\lambda}(2\overline{\mu}+\overline{\kappa})}{\overline{\lambda}+2\overline{\mu}+\overline{\kappa}}h, \quad \overline{\alpha}_2 = \overline{\mu}h, \quad \overline{\alpha}_3 = (\overline{\mu}+\overline{\kappa})h, \quad \overline{\alpha}_4 = (\overline{\mu}+\overline{\kappa})h, \tag{22}$$

$$\overline{\beta}_1 = \overline{\alpha}h - \overline{\mu}\frac{h^3}{12}, \quad \overline{\beta}_2 = \overline{\beta}h - \Lambda\frac{h^3}{12}, \quad \overline{\beta}_3 = \overline{\gamma}h + (2\overline{\mu}+\overline{\kappa}+\Lambda)\frac{h^3}{12}, \quad \beta_4 = \overline{\gamma}h, \tag{23}$$

where h is the plate thickness.

From (22) and (9) it follows the relations for the tangential relaxation functions

$$\frac{P_3}{P_0}\frac{P_1+P_2}{P_1+P_2+P_3}h = \frac{A_1}{A_0}, \quad \frac{P_2}{P_0}h = \frac{A_2}{A_0}, \quad \frac{P_1}{P_0}h = \frac{A_3}{A_0} = \frac{A_4}{A_0}. \tag{24}$$

It is clear that determination of the polynomials $A_k(p)$ from the latter equations can not be performed uniquely. Solving (24) results in

$$A_0 = P_0(P_1+P_2+P_3), \qquad A_1 = P_3(P_1+P_2)h, \tag{25}$$

$$A_2 = P_2(P_1+P_2+P_3)h, \qquad A_3 = A_4 = P_1(P_1+P_2+P_3)h. \tag{26}$$

In a similar way from (23) and (10) we obtain the equations for the bending relaxation functions

$$\frac{B_1}{B_0} = \frac{Q_3}{Q_0}h - \frac{P_2}{P_0}\frac{h^3}{12}, \quad \frac{B_2}{B_0} = \frac{Q_2}{Q_0}h - \frac{P_3(P_1 + P_2)}{P_0(P_1 + P_2 + P_3)}\frac{h^3}{12},$$

$$\frac{B_3}{B_0} = \frac{Q_1}{Q_0}h + \frac{(P_1 + P_2)(P_1 + P_2 + 2P_3)}{P_0(P_1 + P_2 + P_3)}\frac{h^3}{12}, \quad \frac{B_4}{B_0} = \frac{Q_1}{Q_0}h.$$

From this it follows that

$$B_0 = P_0 Q_0 (P_1 + P_2 + P_3),$$

$$B_1 = \left(Q_3 P_0 h - P_2 Q_0 \frac{h^3}{12}\right)(P_1 + P_2 + P_3),$$

$$B_2 = Q_2 P_0 (P_1 + P_2 + P_3)h - P_3 Q_0 (P_1 + P_2)\frac{h^3}{12},$$

$$B_3 = Q_1 P_0 (P_1 + P_2 + P_3)h + Q_0 (P_1 + P_2)(P_1 + P_2 + 2P_3)\frac{h^3}{12},$$

$$B_4 = Q_1 P_0 (P_1 + P_2 + P_3)h.$$

Hence, the differential operators in the constitutive equation for the resultant stress and couple stress tensors $\mathbf{T}$ and $\mathbf{M}$ are more complicated than in the three-dimensional case because the order of derivatives of A_i and B_i is higher than the order of P_k and Q_k, in general. This means that the viscoelastic properties of the two-dimensional structure are more complicated then their three-dimensional counterparts.

Remark. In the case of shells the stress resultant and couple stress tensors depend on $\boldsymbol{\sigma}$ as follows:

$$\overline{\mathbf{T}} = \langle (\mathbf{A} - z\mathbf{B})^{-1} \cdot \overline{\boldsymbol{\sigma}} \rangle, \quad \overline{\mathbf{M}} = \langle (\mathbf{A} - z\mathbf{B})^{-1} \cdot \overline{\boldsymbol{\mu}} \rangle - \langle (\mathbf{A} - z\mathbf{B})^{-1} \cdot z\overline{\boldsymbol{\sigma}} \times \boldsymbol{n} \rangle, \quad (27)$$

where $\mathbf{B} = -\nabla_S \boldsymbol{n}$ is the curvature tensor of the base shell surface, and $\langle \cdot \rangle$ takes a more complex form; see [Lebedev et al. 2010] for details. This means that for a curved surface one should obtain more complicated relaxation properties for two-dimensional theories of shell and plate.

5. Examples

Let us consider the simple case of the viscoelastic micropolar constitutive equations and the corresponding two-dimensional equations.

5.1. *Maxwell model.* For a Maxwell-type model the polynomials in (6) and (7) have the form

$$P_0 = 1 + p_0 p, \qquad P_1, P_2, P_3 = \text{const}, \tag{28}$$
$$Q_0 = 1 + q_0 p, \qquad Q_1, Q_2, Q_3 = \text{const}.$$

For the stress resultant tensor $\mathbf{T}$ we obtain the constitutive equation of Maxwell-type, i.e. A_0 is the linear polynomial only, while the A_1, A_2, A_3, and A_4 are constants. For the couple stress tensor $\mathbf{M}$ we obtain more complicated form of constitutive equations. The simplification is possible if $p_0 = q_0 = 1/\tau$. In this case we use the polynomials

$$B_0 = P_0(P_1 + P_2 + P_3),$$
$$B_1 = \left(Q_3 h - P_2 \frac{h^3}{12}\right)(P_1 + P_2 + P_3),$$
$$B_2 = Q_2(P_1 + P_2 + P_3)h - P_3(P_1 + P_2)\frac{h^3}{12},$$
$$B_3 = Q_1(P_1 + P_2 + P_3)h + (P_1 + P_2)(P_1 + P_2 + 2P_3)\frac{h^3}{12},$$
$$B_4 = Q_1(P_1 + P_2 + P_3)h,$$

where B_0 is a polynomial while B_1, B_2, B_3, and B_4 are constants. Thus, we obtain the Maxwell-type two-dimensional constitutive equations in the case when the relaxation times for the stress and couple stress coincide, i.e., when $P_0 \equiv Q_0$, and we have one relaxation time τ. In the general case we obtain more general constitutive equation for $\mathbf{M}$.

5.2. *Voigt model.* Assume the polynomials in (6) and (7) have the form

$$P_0 = 1, \quad P_1 = P_1^0(1 + p_1 p), \quad P_2 = P_2^0(1 + p_2 p), \quad P_3 = P_3^0(1 + p_3 p),$$
$$Q_0 = 1, \quad Q_1 = Q_1^0(1 + q_1 p), \quad Q_2 = Q_2^0(1 + q_2 p), \quad Q_3 = Q_3^0(1 + q_3 p). \tag{29}$$

Here the corresponding polynomials A_i, $i = 1, 2, 3, 4$, are linear functions of p. This means that for $\mathbf{T}$ we have the viscoelastic equations which are similar to the standard viscoelastic model [Christensen 1971]. For $\mathbf{M}$ we again obtain a more complicated model because in the general case. B_0 is a linear function while B_1, B_2, B_3, and B_4 are quadratic functions of p.

A simplification of the two-dimensional constitutive equation for $\mathbf{M}$ is possible if we assume $p_1 = p_2 = p_3 = p_4 = q_1 = q_2 = q_3 = q_4 = 1/\tau$. In this case we have

$$B_0 = (P_1^0 + P_2^0 + P_3^0),$$
$$B_1 = \left(Q_3^0 h - P_2^0 \frac{h^3}{12}\right)(P_1^0 + P_2^0 + P_3^0)\left(1 + \frac{p}{\tau}\right),$$
$$B_2 = \left(Q_2^0(P_1^0 + P_2^0 + P_3^0)h - P_3^0(P_1^0 + P_2^0)\frac{h^3}{12}\right)\left(1 + \frac{p}{\tau}\right),$$
$$B_3 = \left(Q_1^0(P_1^0 + P_2^0 + P_3^0)h + (P_1^0 + P_2^0)(P_1^0 + P_2^0 + 2P_3^0)\frac{h^3}{12}\right)\left(1 + \frac{p}{\tau}\right),$$
$$B_4 = Q_1^0(P_1^0 + P_2^0 + P_3^0)h\left(1 + \frac{p}{\tau}\right).$$

As for the Maxwell-type model, we conclude that the two-dimensional constitutive equations of Voigt-type exists as a special case of the three-dimensional viscoelastic behavior, i.e., with the same relaxation parameter τ.

Conclusion

We have discussed the two-dimensional constitutive equations for resultant force stress and couple stress tensors derived from the constitutive equations of three-dimensional viscoelastic Cosserat (micropolar) continuum. The presented results demonstrate how the viscoelastic properties of three-dimensional continuum inherit in the constitutive equations for plates and shells. Within the framework of the linear micropolar viscoelasticity with constitutive equations of differential type we show that 2D relaxation functions of shells have move complicated structures then the relaxation function of the bulk material, in general. In particular, even for homogeneous plates and shells the spectrum of relaxation time do not coincide with the spectrum of the bulk material. For inhomogeneous plates and shells the spectrum may depend also on the structure of the shell in the thickness direction and its curvature in the case of shells.

Appendix: Through-the-thickness integration

Following [Altenbach and Eremeyev 2009] we present more details on the through-the-thickness procedure used for the derivation of 2D governing equations of the micropolar theory of plates. Let $V = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in \mathcal{M} \subset \mathbb{R}^2, z \in [-h/2, h/2]\}$ be the volume of a plate-like body. We denote the boundary of the plate-like body as $S = S^{\nu} \bigcup S^+ \bigcup S^-$, where $S^{\pm} = \{(x, y, z) : (x, y) \in \mathcal{M}, z = \pm h/2\}$ are the plate faces and $S^{\nu} = \{(x, y, z) : (x, y) \in \mathcal{C} \equiv \partial\mathcal{M}, z \in [-h/2, h/2]\}$ is the lateral surface. We consider the following boundary conditions at $S^{\pm}$

$$n^{\pm} \cdot \sigma = t^{\pm}, \quad n^{\pm} \cdot \mu = m^{\pm}, \tag{30}$$

where $t^{\pm}, m^{\pm}$ are given vector functions, $n^{\pm} = \pm i_3$, and $i_3 = i_1 \times i_2$.

Integrating $(1)_1$ over $[-h/2, h/2]$ and introducing the notations

$$q = \langle \rho f \rangle + t^+ + t^- \tag{31}$$

we result in $(12)_1$ with $\mathbf{T} = \langle \mathbf{A} \cdot \sigma \rangle$.

For the derivation of $(12)_2$ we cross-multiply $(1)_2$ by $z i_3$ from the left and again integrate the result through the thickness. Finally, we obtain $(12)_2$ with

$$c = \langle \rho \ell \rangle + m^+ + m^- + i_3 \times \langle \rho z f \rangle + \frac{h}{2} i_3 \times (t^+ - t^-)$$

and $\mathbf{M} = \langle \mathbf{A} \cdot \mu \rangle - \langle \mathbf{A} \cdot z\sigma \times i_3 \rangle$.

In the case of linear theory the through-the-thickness integration can be applied to the 3D constitutive relations with the linear approximation of displacements and rotations

$$\boldsymbol{u}(x, y, z) = \boldsymbol{v}(x, y) - z\boldsymbol{\phi}(x, y), \quad \boldsymbol{\vartheta} = \boldsymbol{\phi}(x, y) \times \boldsymbol{i}_3 + \vartheta_3(x, y)\boldsymbol{i}_3, \qquad (32)$$

where $\boldsymbol{\phi} \cdot \boldsymbol{i}_3 = 0$. The approximation (32) is consistent with 2D equilibrium equations as well as with averaging through the thickness. In particular, 2D fields of translations and rotations can be interpreted as

$$\boldsymbol{v} = \frac{1}{h}\langle \boldsymbol{u} \rangle, \quad \boldsymbol{\theta} = \boldsymbol{\vartheta}.$$

Other possible variants of 3D to 2D reduction within micropolar elasticity are discussed in [Altenbach et al. 2009; Altenbach and Eremeyev 2009; 2010; Chróścielewski et al. 2011; Sargsyan 2011; Sargsyan and Sargsyan 2014; Steinberg and Kvasov 2013; 2015; Zubov 2009].

Acknowledgements

References

[Altenbach and Eremeyev 2009] H. Altenbach and V. A. Eremeyev, "On the linear theory of micropolar plates", *ZAMM Z. Angew. Math. Mech.* **89**:4 (2009), 242–256.

[Altenbach and Eremeyev 2010] H. Altenbach and V. A. Eremeyev, "On the theories of plates based on the Cosserat approach", pp. 27–35 in *Mechanics of generalized continua: one hundred years after the Cosserats*, edited by G. A. Maugin and A. V. Metrikine, Adv. Mech. Math. **21**, Springer, New York, 2010.

[Altenbach et al. 2009] J. Altenbach, H. Altenbach, and V. A. Eremeyev, "On generalized Cosserat-type theories of plates and shells: a short review and bibliography", *Arch. Appl. Mech.* **80**:1 (2009), 73–92.

[Bîrsan and Neff 2013] M. Bîrsan and P. Neff, "Existence theorems in the geometrically non-linear 6-parameter theory of elastic plates", *J. Elasticity* **112**:2 (2013), 185–198.

[Bîrsan and Neff 2014] M. Bîrsan and P. Neff, "Existence of minimizers in the geometrically non-linear 6-parameter resultant shell theory with drilling rotations", *Math. Mech. Solids* **19**:4 (2014), 376–397.

[Christensen 1971] R. M. Christensen, *Theory of viscoelasticity: an introduction*, Academic Press, New York, 1971.

[Chróścielewski et al. 2004] J. Chróścielewski, J. Makowski, and W. Pietraszkiewicz, *Statyka i dynamika powłok wielopłatowych: nieliniowa teoria i metoda elementów skończonych*, Wydawnictwo IPPT PAN, Warsaw, 2004.

[Chróścielewski et al. 2010] J. Chróścielewski, W. Pietraszkiewicz, and W. Witkowski, "On shear correction factors in the non-linear theory of elastic shells", *Int. J. Solids Struct.* **47** (2010), 3537–3545.

[Chróścielewski et al. 2011] J. Chróścielewski, I. Kreja, A. Sabik, and W. Witkowski, "Modeling of composite shells in 6-parameter nonlinear theory with drilling degree of freedom", *Mech. Adv. Mater. Struct.* **18** (2011), 403–419.

[Diebels and Steeb 2003] S. Diebels and H. Steeb, "Stress and couple stress in foams", *Comput. Mater. Sci.* **28** (2003), 714–722.

[Eremeyev and Lebedev 2011] V. A. Eremeyev and L. P. Lebedev, "Existence theorems in the linear theory of micropolar shells", *Z. Angew. Math. Mech.* **91**:6 (2011), 468–476.

[Eremeyev and Zubov 2008] V. A. Eremeyev and L. M. Zubov, *Mekhanika uprugikh obolochek*, Nauka, Moscow, 2008.

[Eremeyev et al. 2013] V. A. Eremeyev, L. P. Lebedev, and H. Altenbach, *Foundations of micropolar mechanics*, Springer, Heidelberg, 2013.

[Ericksen and Truesdell 1958] J. L. Ericksen and C. Truesdell, "Exact theory of stress and strain in rods and shells", *Arch. Ration. Mech. An.* **1** (1958), 295–323.

[Eringen 1999] A. C. Eringen, *Microcontinuum field theories, I: Foundations and solids*, Springer, New York, 1999.

[Lakes 1986] R. S. Lakes, "Experimental microelasticity of two porous solids", *Int. J. Solids Struct.* **22** (1986), 55–63.

[Lebedev et al. 2010] L. P. Lebedev, M. J. Cloud, and V. A. Eremeyev, *Tensor analysis with applications in mechanics*, World Scientific, Hackensack, NJ, 2010.

[Libai and Simmonds 1998] A. Libai and J. G. Simmonds, *The nonlinear theory of elastic shells*, 2nd ed., Cambridge University Press, 1998.

[Reissner 1977] E. Reissner, "A note on generating generalized two-dimensional plate and shell theories", *Z. Angew. Math. Phys.* **28** (1977), 633–642.

[Rubin 2000] M. B. Rubin, *Cosserat theories: shells, rods and points*, Kluwer, Dordrecht, 2000.

[Sargsyan 2011] S. H. Sargsyan, "The general dynamic theory of micropolar elastic thin shells", *Dokl. Phys.* **56**:1 (2011), 39–42.

[Sargsyan and Sargsyan 2014] A. H. Sargsyan and S. H. Sargsyan, "Dynamic model of micropolar elastic thin plates with independent fields of displacements and rotations", *J. Sound Vib.* **333**:18 (2014), 4354–4375.

[Steinberg and Kvasov 2013] L. Steinberg and R. Kvasov, "Enhanced mathematical model for Cosserat plate bending", *Thin-Walled Struct.* **63** (2013), 51–62.

[Steinberg and Kvasov 2015] L. Steinberg and R. Kvasov, "Analytical modeling of vibration of micropolar plates", *Appl. Math.* (*SCIRP*) **6**:5 (2015), 817–836.

[Wilson 1901] E. B. Wilson, *Vector analysis: a textbook for the use of students of mathematics and physics, founded upon the lectures of J. Willard Gibbs*, Yale University Press, New Haven, 1901. Reprinted by Dover , New York, 1960.

[Zubov 2009] L. M. Zubov, "Micropolar-shell equilibrium equations", *Dokl. Phys.* **54**:6 (2009), 290–293.

HOLM ALTENBACH: holm.altenbach@ovgu.de
Chair of Engineering Mechanics, Institute of Mechanics, Faculty of Mechanical Engineering, Otto-von-Guericke-Universität, Universitätsplatz 1, 39106 Magdeburg, Germany

VICTOR A. EREMEYEV: eremeyev.victor@gmail.com
Department of Applied Mechanics and Robotics, Rzeszow University of Technology, al. Powstańców Warszawy, 8, 35-959, Rzeszów, Poland
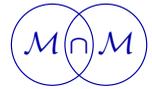
IDENTIFICATION OF HIGHER-ORDER
ELASTIC CONSTANTS FOR GRAIN ASSEMBLIES
BASED UPON GRANULAR MICROMECHANICS

ANIL MISRA AND PAYAM POORSOLHJOUY

Macroscale behavior of granular media is characterized by the significant effects of grain-pair interactions and the microstructure of each grain neighborhood. From a continuum viewpoint, granular materials may be modeled as micromorphic media to account for their complex grain-scale (microscale) kinematics. To this end we express the grain displacement in terms of the neighboring grain displacements utilizing the Taylor series expansion. The introduced gradients in the Taylor series are identified in terms of the macroscale deformation measures introduced in microstructural elasticity and micromorphic mechanics. As a result, a continuum model of the granular media is derived enriched by nonclassical terms, including terms that model grain displacement fluctuations and higher gradients of displacements. In the derived model, the continuum stiffness tensors are obtained in terms of grain-pair stiffness coefficients and fabric parameters defining the geometry of grains and their contacts. To identify the elastic constants of the enhanced continuum model, we perform numerical experiments on grain assemblies using discrete simulations subjected to relevant boundary conditions. The need for additional macroscale deformation measures for the continuum modeling of granular materials becomes evident in this identification process. The obtained elastic constants are then used to determine the microscale (or grain-pair) stiffness coefficients applicable to the continuum model. These grain-scale stiffness coefficients are found to be affected by the heterogeneity of microstructure.

1. Introduction

Among the various forms of granular materials, the grain assemblies formed of grains in noncohesive contacts more strongly exhibit the microstructural and micromechanical effects at the macroscale behavior. For a macroscale description of these materials, therefore, approaches are needed that can readily represent
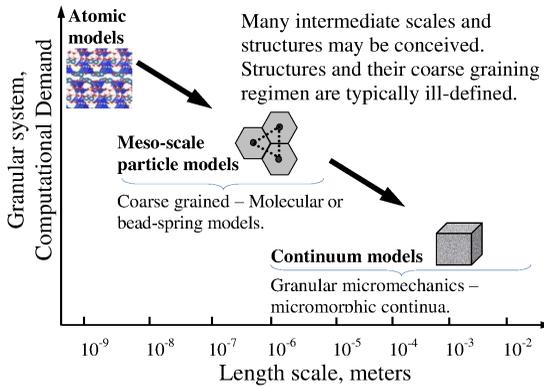
Figure 1. Schematic representation of modeling length scales and the corresponding computational demand.

the effects of grain-pair interactions. Discrete models based upon various coarse-graining schemes as schematically depicted in Figure 1 are not desirable for many problems in science and engineering that deal with structures that may contain large numbers of grains ($> 10^6$). These models are not only computationally expensive, but require local constitutive laws for grain-pair interactions and simulated microstructures, which are often not readily conceivable. Multiscale models that seek a hybrid discrete-continuum approach also suffer from reliance on information at various scales that is not readily available or easily conceived for complex material systems. Nonclassical continuum models are needed that can capture the effects of microscale mechanisms. Indeed, the necessity of modeling microscale mechanisms within the rubric of continuum mechanics is made clear in pioneering works such as [Cosserat and Cosserat 1909; Eringen 1999; Germain 1973; Green and Rivlin 1964; Mindlin 1964; Toupin 1964]. The seminal developments of continuum mechanics are known to proceed from some micromechanical conception of deformable materials, as seen from the works of Piola [Auffray et al. 2015; dell'Isola et al. 2014], Navier [1827] and Cauchy [1826–1830].

An important aspect of deformation of grain packing is nonaffine particle movement when a volume element of these materials is subjected to uniform boundary conditions. This aspect has been observed experimentally [Misra 1998; Misra and Jiang 1997], through many numerical simulations (see among others [Peters and Walizer 2013]), and in models of these materials within the classical continuum mechanics framework [Jenkins et al. 2005; Misra and Chang 1993; Trentadue 2001]. To model the effects of fluctuations in grain motions (termed as micro-deformations or inner motions) it is advantageous to enrich the classical continuum approach [Maugin 2014]. To that end we combine the granular micromechanics paradigm, which offers a robust methodology for developing continuum models of granular material systems by incorporating microscale effects [Misra and Singh 2014;

2015; Misra and Poorsolhjouy 2015c], with the identification of grain (microscale) motions in terms of the Mindlin–Eringen macroscale deformation measures introduced in microstructural elasticity [Mindlin 1964] and micromorphic mechanics [Eringen 1999]. As a result, an enhanced continuum model of the granular media is obtained, enriched by nonclassical terms that model grain displacement fluctuations and higher gradients of displacements [Misra and Poorsolhjouy 2015b; 2015a]. In addition, expressions are derived for effective macroscopic constitutive coefficients corresponding to (1) the macroscale displacement gradient, (2) the fluctuations in displacement gradient which is related to the microscale displacement gradient, and (3) the macroscale gradient of microdisplacement gradient or the second gradient of displacement. In this paper, we demonstrate a method for identifying the elastic constants of the enhanced continuum model based upon numerical experiments on grain assemblies using discrete simulations. Two types of grain assemblies are analyzed: (1) regular closed-packed assembly of equal-diameter disks with vacancy type defects, and (2) irregular assembly of disks of three different sizes. These assemblies are treated as 2D granular media for modeling purposes. Discrete simulations are performed using unit cells of these grain assemblies to obtain deformation energies under prescribed displacement boundary conditions compatible with uniform continuum kinematic measures. These deformation energies are then analyzed to identify the continuum elastic constants. The need for additional macroscale deformation measures for the continuum modeling of granular materials becomes evident in this identification process. The obtained elastic constants are used to determine the microscale (or grain-pair) stiffness coefficients applicable to the continuum model. These grain-scale stiffness coefficients are found to be affected by the heterogeneity of microstructure, indicating that they do not represent stiffness of an isolated grain-pair; rather they represent a collective stiffness behavior of the grain-pair and its neighbors.

2. Enhanced continuum model of granular media

2.1. *Micro-macro kinematic identification.* We consider a unit cell representative of the granular assembly. In a continuum model, the unit cell plays the role of a material point P, as depicted in Figure 2, where the macroscale coordinate system $\boldsymbol{x}$ only distinguishes different material points. In addition, a coordinate system $\boldsymbol{x}'$ is attached to the material point, P, at the barycenter of the unit cell with coordinate axes parallel to those of the macroscale coordinate system $\boldsymbol{x}$. Following the methodology of microstructural elasticity described by Mindlin [1964], we define the micro-deformation gradient, $\psi_{ij}(x_k) = \phi_{i,j}(x_k)$, which is decomposed into a part representing the average or macroscale displacement gradient, $\bar{\phi}_{i,j}$, and a second part, γ_{ij}, representing the gradient of the displacement fluctuations with
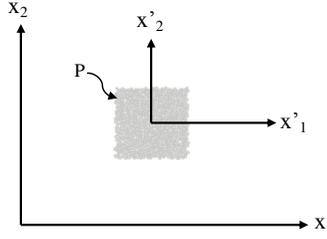
Figure 2. Representation of material point and the global coordinate system, $\boldsymbol{x}$, and the local coordinate system, $\boldsymbol{x}'$, located at the material point's center of mass.

respect to $\boldsymbol{x}'$, as

$$\psi_{ij} \triangleq \phi_{i,j} = \bar{\phi}_{i,j} - \gamma_{ij}. \tag{1}$$

From the viewpoint of identifying constitutive relationships, we consider the continuum to be subjected to a linear macroscale displacement field. In this case, the macroscale displacement gradient, $\bar{\phi}_{i,j}$, is independent of coordinates x_k. Further, the grain displacement fluctuations are assumed to be linear functions of the local coordinate system $\boldsymbol{x}'$, leading to a micromorphic theory of degree 1, in the terminology introduced by Germain [1973]. This implies that the second-rank tensor γ_{ij} is independent of the local coordinates (it is homogeneous within the material point) and is a function only of x_k. Throughout the paper, the subscripts follow tensor notation and the summation convention over repeated subscripts is implied unless explicitly noted otherwise.

We now consider the relative displacement of two contacting neighbor grains, n and p, within the unit cell of the granular media. Displacement of grain p is written using a Taylor series expansion of the displacement of the neighbor grain, n, with terms up to second order, as

$$\delta_i^{np} = \phi_i^p - \phi_i^n = \phi_{i,j}^n l_j + \tfrac{1}{2}\phi_{i,jk}^n l_j l_k + \cdots, \tag{2}$$

where ϕ_i is the displacement of grain centroids, and l_j is the vector joining the centroids of the two grains, n and p (denoted by the superscripts). Using (1) and (2), the relative displacement of two neighbor grains, n and p, is derived as

$$\delta_i^{np} = (\bar{\phi}_{i,j} - \gamma_{ij})l_j + \tfrac{1}{2}\phi_{i,jk}l_j l_k = \bar{\phi}_{i,j}l_j - \gamma_{ij}l_j + \tfrac{1}{2}\phi_{i,jk}l_j l_k = \delta_i^M - \delta_i^m + \delta_i^g. \tag{3}$$

As seen from (3), the intergranular relative displacements between two interacting (contacting) grains is decomposed into three terms:

(1) δ_i^M, due to the macroscale displacement gradient, $\bar{\phi}_{i,j}$;

(2) δ_i^m, due to the gradients of the fluctuation in grain displacement, $\gamma_{ij}(x)$; and

(3) δ_i^g, due to the second gradient term, $\phi_{i,jk}$, which is same as the gradient of the relative displacements, $\gamma_{ij,k}$.

The three relative displacement terms, denoted by the superscripts M, m and g, are

$$\delta_i^M = \bar{\phi}_{i,j} l_j; \quad \delta_i^m = \gamma_{ij} l_j; \quad \delta_i^g = \tfrac{1}{2}\phi_{i,jk} l_j l_k = \phi_{i,jk} J_{jk} = \psi_{ij,k} J_{jk}, \quad (4)$$

and the tensor $J_{ij} = l_i l_j / 2$ represents a moment tensor introduced here for simplifying further derivations.

The relative rotation of grains within the granular assembly can be related to the rotation field within the material point, defined as the curl of displacement field [Misra and Poorsolhjouy 2015b]. Thus, applying Taylor series expansion, the relative rotation of two neighbor grains, n and p, denoted as $\boldsymbol{\theta}$, is obtained as

$$\theta_i^{np} = \kappa_i^p - \kappa_i^n = \kappa_{i,p}^n l_p = (e_{ijk}\phi_{k,j})_{,p} l_p = e_{ijk}\phi_{k,jp} l_p. \quad (5)$$

The intergranular relative rotation between two interacting grains is, thus, related to the second gradient term, $\phi_{i,jk}$. That grains undergo relative rotations is also known from measurements of kinematic fields in grain assembles [Misra 1998; Misra and Jiang 1997].

2.2. Stress and force conjugates to macro-micro kinematic measures. For further development, we retain only the symmetric part of the macroscale displacement gradient tensor, $\bar{\phi}_{(i,j)}$, as the classical small-deformation strain tensor, and express the macroscale deformation energy density of the granular continua as a function of the continuum kinematic measures as $W = W(\bar{\phi}_{(i,j)}, \gamma_{ij}, \phi_{i,jk})$. The macroscale stress components conjugate to the kinematic measures are then defined as

$$\tau_{ij} = \frac{\partial W}{\partial \bar{\phi}_{(i,j)}} = \frac{\partial W}{\partial \epsilon_{ij}}, \quad \sigma_{ij} = \frac{\partial W}{\partial \gamma_{ij}}, \quad \mu_{ijk} = \frac{\partial W}{\partial \gamma_{ij,k}}, \quad (6)$$

where τ_{ij}, σ_{ij}, and μ_{ijk} are Cauchy stress, relative stress, and double stress, respectively. We note that the macroscopic strain energy density function can be obtained as the volume average of the deformation energies of the grain-pair interactions, written as

$$W = \frac{1}{V} \sum_\alpha W^\alpha(\delta_i^M, \delta_i^m, \delta_i^g, \theta_i^u), \quad (7)$$

where the superscript α denotes the α-th grain-pair interaction, and it has been intentionally dropped from the kinematic measures to simplify the equations. The forces and moment conjugate to the microscale kinematic measures are defined as

$$\frac{\partial W^\alpha}{\partial \delta_i^{\alpha\xi}} = f_i^{\alpha\xi}, \quad \text{where } \xi = M, m, g, \quad \text{and} \quad \frac{\partial W^\alpha}{\partial \theta_i^{\alpha u}} = m_i^{\alpha u}. \quad (8)$$

Now, substituting (8) and (7) into (6) and using (4) and (5), respectively, macroscopic stress tensors conjugate to the macroscale kinematic measures are obtained as

$$3\tau_{ij} = \frac{\partial W}{\partial \epsilon_{ij}} = \frac{1}{V}\sum_{\alpha}\frac{\partial W^{\alpha}}{\partial \delta_k^M}\frac{\partial \delta_k^M}{\partial \epsilon_{ij}} = \frac{1}{V}\sum_{\alpha}f_i^{M\alpha}l_j^{\alpha}, \tag{9}$$

$$\sigma_{ij} = \frac{\partial W}{\partial \gamma_{ij}} = \frac{1}{V}\sum_{\alpha}\frac{\partial W^{\alpha}}{\partial \delta_k^m}\frac{\partial \delta_k^m}{\partial \gamma_{ij}} = \frac{1}{V}\sum_{\alpha}f_i^{m\alpha}l_j^{\alpha}, \tag{10}$$

$$\mu_{ijk} = \frac{\partial W}{\partial \phi_{i,jk}} = \frac{1}{V}\sum_{\alpha}\left(\frac{\partial W^{\alpha}}{\partial \delta_l^g}\frac{\partial \delta_l^g}{\partial \phi_{i,jk}} + \frac{\partial W^{\alpha}}{\partial \theta_l^u}\frac{\partial \theta_l^u}{\partial \phi_{i,jk}}\right)$$

$$= \frac{1}{V}\left(\sum_{\alpha}f_i^{g\alpha}J_{jk}^{\alpha} + \sum_{\alpha}m_l^{u\alpha}e_{jil}l_k^{\alpha}\right). \tag{11}$$

Thus, the stress tensor conjugates are related to the microscale force measures and moments. Similar expressions have been reported in literature [Chen and Lee 2003] based upon the generalization of the virial theorem [Ganghoffer 2010]. However, it is notable that, in the stress tensor expressions presented here, separate force measures conjugate to the grain-pair (microscale) counterparts of the continuum deformation measures are defined, which is clearly different from those based upon the virial theorem.

2.3. *Microscale and macroscale constitutive equations.* For linear isotropic elasticity, the following quadratic form of W^{α} is formulated:

$$W^{\alpha} = \frac{1}{2}\left[\sum_{\xi}K_{ij}^{\alpha\xi}\delta_i^{\alpha\xi}\delta_j^{\alpha\xi} + G_{ij}^{\alpha u}\theta_i^{\alpha u}\theta_j^{\alpha u}\right], \quad \text{where } \xi = M, m, g, \tag{12}$$

where K and G represent the grain-pair stiffness parameters for forces and moments, respectively. Thus, the microscale constitutive equations that link the microscale kinematics measures to their conjugate force and moment measures take the form

$$f_i^{\alpha\xi} = K_{ij}^{\alpha\xi}\delta_j^{\alpha\xi}, \quad \text{where } \xi = M, m, g,$$
$$m_i^{\alpha u} = G_{ij}^{\alpha u}\theta_j^{\alpha u}. \tag{13}$$

The introduced grain-pair stiffness parameters define the force conjugates associated with different microscale kinematic measures that contribute to intergranular relative displacements and rotations. These stiffness parameters do not represent the stiffness of two isolated interacting grains. In the derived model we have introduced four types of intergranular stiffness parameters, namely, the average, the fluctuation, the second gradient and the rotational, distinguished by their superscripts M, m, g and u, respectively. We note that in (12) and (13) the terms that cross-link the different microscale kinematic measures have been ignored for

simplicity. The consequence of this assumption will be that uncoupled constitutive relationships will be obtained for the continuum model. Further, it should be noted that the microscale stiffness coefficients in (13) represent intergranular stiffness measures in a statistical sense. Since the total relative displacement between two neighbor grains was previously decomposed into different terms, representing the average and fluctuation displacement gradients and also second gradient terms, the stiffness coefficients corresponding to them will all contribute, in a statistical sense, to form the overall response of the interaction to the grains' relative displacements. Now, substituting the microscale constitutive equations, (13), into (9)–(11), and noting that the macroscale strain measures are constant throughout the material point, macroscale constitutive equations are derived as

$$\tau_{ij} = \frac{1}{V}\sum_\alpha f_i^M l_j^\alpha = \frac{1}{V}\sum_\alpha K_{ik}^M \delta_k^M l_j^\alpha = \left(\frac{1}{V}\sum_\alpha K_{ik}^M l_l^\alpha l_j^\alpha\right)\epsilon_{kl} = C_{ijkl}^M \epsilon_{kl}, \quad (14)$$

$$\sigma_{ij} = \frac{1}{V}\sum_\alpha f_i^m l_j^\alpha = \frac{1}{V}\sum_\alpha K_{ik}^m \delta_k^\alpha l_j^\alpha = \left(\frac{1}{V}\sum_\alpha K_{ik}^m l_l^\alpha l_j^\alpha\right)\gamma_{kl} = C_{ijkl}^m \gamma_{kl}, \quad (15)$$

$$\mu_{ijk} = \frac{1}{V}\sum_\alpha (f_i^g J_{jk}^\alpha + m_l^u e_{jil} l_k)$$

$$= \left(\frac{1}{V}\sum_\alpha K_{il}^g J_{mn}^\alpha J_{jk}^\alpha + \frac{1}{V}\sum_\alpha G_{pq}^u e_{mlq} e_{jip} l_k l_n\right)\phi_{l,mn}$$

$$= (A_{ijklmn}^g + A_{ijklmn}^u)\phi_{l,mn} = A_{ijklmn}\phi_{l,mn}. \quad (16)$$

Here it should be noted that the summations in (14)–(16) are performed over all grain-pair interactions inside the material point. The quantities inside the summation (stiffness coefficients and geometric parameters l_i and J_{ij}) are all different for every single grain-pair interaction. It is, however, possible to conceive of average value for all these quantities for grain-pair interactions in any given orientation. Using these average values, the summation over all grain interactions can be changed to integration over all generic orientations. To this end we define N_p as the number density of grain-pair interactions divided by the volume of the unit cell, and ξ as the normalized directional density distribution function of contacts in different orientations within the unit cell. Having a properly defined ξ will enable the method to model materials with different levels of anisotropy. For modeling isotropic materials or particle assemblies with random distribution of grains in different orientations, the distribution function should take a constant value independent of the direction. So, for isotropic materials (or for randomly distributed grain assemblies), the density distribution function in 2D and 3D domains will be

$$\xi(\theta) = \frac{1}{2\pi} \quad \Longrightarrow \quad \int_\theta \xi \, d\theta = \frac{1}{2\pi}2\pi = 1 \quad \text{(for 2D)} \quad (17a)$$

and

$$\xi(\theta, \phi) = \frac{1}{4\pi} \quad \Longrightarrow \quad \int_\theta \int_\phi \xi(\sin\theta \, d\theta \, d\phi) = \frac{1}{4\pi}(2\pi)(2) = 1 \quad \text{(for 3D)}. \quad (17b)$$

In (17a), θ is the polar angle of the 2D polar coordinate system, while in (17b), θ and ϕ are the polar angles measured from the vertical axis and the azimuth angle in the 3D spherical coordinate system, respectively.

3. Identification of elastic constants using discrete simulations

Evidently, the enhanced continuum model of granular media derived in Section 2 can be characterized by either microscale or macroscale elastic constants. To demonstrate the applicability of the derived model, we identify these constants for specific grain assemblies using discrete simulations.

3.1. *Discrete simulation methodology.* For any grain within the assembly, say the n-th grain, the variation of the total potential energy can be written as

$$\delta W^n = \sum_\alpha \delta W^{n\alpha}(\delta_i^{n\alpha}, \theta_i^{n\alpha}) = \sum_\alpha f_i^{n\alpha} \delta \delta_i^{n\alpha} + \sum_\alpha (m_i^{n\alpha} + e_{ijk} f_j^{n\alpha} r_k^n) \delta \theta_i^{n\alpha}, \quad (18)$$

where the summation over α denotes summation over all grains which have interaction with the n-th grain. In (18), $f^{n\alpha}$, $m^{n\alpha}$, $\delta^{n\alpha}$, and $\theta^{n\alpha}$ are the grain-pair force, moment, relative displacement, and relative rotation for grains n and α, respectively, related through the grain-scale constitutive equations for an isolated grain-pair as

$$f_i^{n\alpha} = \bar{K}_{ij}^\alpha \delta_j^{n\alpha}, \quad \text{where } \bar{K}_{ij}^{n\alpha} = \bar{K}_n^\alpha n_i^\alpha n_j^\alpha + \bar{K}_s^\alpha s_i^\alpha s_j^\alpha + \bar{K}_{ns}^\alpha (n_i^\alpha s_j^\alpha + s_i^\alpha n_j^\alpha),$$
$$m^{n\alpha} = \bar{G}^\alpha \theta^{n\alpha}, \quad (19)$$

where the isolated grain-pair stiffness coefficients for 2D contact of disks in the normal, tangential, and normal-tangential coupling are taken to be $\bar{K}_n$, $\bar{K}_s$, and $\bar{K}_{ns}$, respectively, and the grain-pair moment stiffness is taken to be $\bar{G}$.

Denoting the total force exerted on the n-th grain by F_i and the total moment by M_i, the variation of the total potential energy is written as

$$\delta W^n = F_i^n \delta \phi_i^n + M_i^n \delta \kappa_i^n. \quad (20)$$

Further, the variation of the relative kinematic measures can be derived based on the variation of displacement and rotation fields as

$$\delta \delta_i^{n\alpha} = \delta(\phi_i^\alpha - \phi_i^n) = -\delta \phi_i^n,$$
$$\delta \theta_i^{n\alpha} = \delta(\kappa_i^\alpha - \kappa_i^n) = -\delta \kappa_i^n. \quad (21)$$

Setting (18) and (20) equal and substituting (21), the following balance equations for grain, n, is found:

$$F_i^n = -\sum_\alpha f_i^{n\alpha},$$

$$M_i^n = -\sum_\alpha (m_i^{n\alpha} + e_{ijk} f_j^{n\alpha} r_k^n). \tag{22}$$

Combining (19) and (22), the total grain force and moment can be related to the grain motions. These expressions can be assembled to form an overall force-displacement equation governing the behavior of the assembly as a whole, written as [Chang and Misra 1989]

$$\{F\}_{3M \times 1} = [S]_{3M \times 3M} \{u\}_{3M \times 1}, \tag{23}$$

where M denotes the total number of grains within the whole assembly. For each grain there are three kinematic measures (two displacements and one rotation), resulting in a total number of $3M$ kinematic variables represented in u. Also, each grain has three force measures (two forces and one moment), resulting in a total number of $3M$ force variables represented in F. So the overall number of variables in (23) is $6M$, knowing any $3M$ of which the other $3M$ can be derived by solving the system of equations represented in matrix form in (23).

3.2. Methodology for identification of constitutive coefficients. Using (6), the macroscale strain energy can be written as

$$W = C_{ijkl}^M \epsilon_{ij} \epsilon_{kl} + C_{ijkl}^m \gamma_{ij} \gamma_{kl} + A_{ijklmn} \psi_{ij,k} \psi_{lm,n}. \tag{24}$$

For finding the components of the stiffness tensors, numerical experiments are performed with 2D grain assembly using the discrete simulation method. These numerical experiments are performed in a manner akin to physical experiments by applying boundary conditions on the grain assemblies. Thus, to obtain the components of the stiffness tensor, C_{ijkl}^M, we purely apply on the boundary of the grain assembly a specified macroscale displacement gradient while constraining the fluctuation displacement gradient, γ_{ij}, and the second gradient term, $\psi_{ij,k}$, to be zero. In order to achieve this type of displacement at the boundary, boundary grains and their immediate neighbors are displaced by the amount compatible with the specified macroscale displacement gradient. In this manner, the fluctuations in the displacement gradient and the second gradient of displacement are ensured to be zero. For the assembly loaded in this manner, the strain energy will be

$$W|_{(\gamma_{ij}=0, \psi_{ij,k}=0)} = C_{ijkl}^M \epsilon_{kl} \epsilon_{ij}. \tag{25}$$

Thus, by performing a sufficient number of simulations with specified combinations of the strain components, the elastic constants are identified.

For finding the components of C_{ijkl}^m, we purely apply fluctuations in displacement gradient while keeping the macroscale displacement gradient and the second gradient of displacement to be zero. For this type of boundary condition, the boundary grains are kept stationary while their immediate neighbors are moved according to a fluctuation in displacement gradient using (3). By loading the assembly in this manner, the strain energy of the assembly will be

$$W|_{(\epsilon_{ij}=0,\,\psi_{ij,k}=0)} = C_{ijkl}^m \gamma_{ij}\gamma_{kl}. \tag{26}$$

Applying a sufficient number of combinations of fluctuation strain components as for the case of macrostrain, the components of the tensor C_{ijkl}^m are derived. Finally, for finding the components of the sixth-rank stiffness tensors, second gradients of displacement fluctuations are applied on the assembly while keeping the average and fluctuations in displacement gradient to be equal to zero. For this purpose the boundary grains are kept stationary while their immediate neighbors are moved in consistence with the second gradient in the displacement fluctuations. Movements of the neighbors of boundary grains are derived using (3) with the desired value of $\psi_{ij,k}$. In this loading scheme the internal strain energy will be derived as

$$W|_{(\epsilon_{ij}=0,\,\gamma_{ij}=0)} = A_{ijklmn}\psi_{ij,k}\psi_{lm,n}. \tag{27}$$

For a 2D granular system, the sixth-rank stiffness tensor A_{ijklmn} can be written as an 8×8 matrix with 36 independent components. To find all these components, second gradient of displacement in 36 different combinations need to be applied to the material. Eight of these combinations are indicated by the eight different components of ψ_{ijk} and the 26 remaining ones are composed of their combinations. Finding the energy for each one of these cases, using (18) and writing (27), result in a system of 36 equations and 36 unknowns that, when solved, will yield all components of the second gradient stiffness tensor.

4. Results

Two types of grain assemblies are analyzed: (1) regular closed-packed assembly of equal-diameter disks (monodisperse) with vacancy type defects, and (2) irregular assembly of three different diameter disks (tridisperse). These assemblies are treated as 2D granular media for modeling purposes.

4.1. *Regular monodispersed grain assembly.* Regular hexagonal closed-packed assemblies from cylindrical particles (assembly A, seen in Figure 3, left) have been analyzed. Further, defected assemblies are also made by removing grains from the assembly once every four rows (assembly B, seen in Figure 3, middle) and once
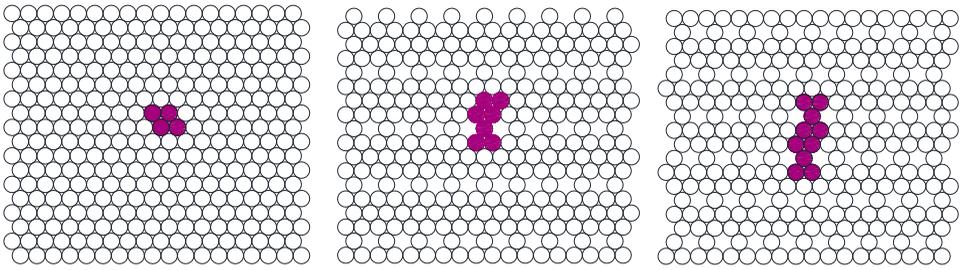
Figure 3. Three regular assemblies. Left: hexagonal closed-pack (assembly A). Middle and right: defected assemblies (assemblies B and C, respectively).

every three rows (assembly C, seen in Figure 3, right). Thus, Assembly A is the densest of the three, B is intermediate, and C is the loosest.

Assembly A has planes of elastic symmetry every 30 degrees. Applying the effects of elastic symmetries about horizontal and vertical planes shows that the material behaves as an orthotropic material in 2D with four independent constants in its first gradient stiffness tensor. Considering elastic symmetry about the plane whose normal vector makes a 150° angle with the positive x-axis results in a stiffness tensor with only two independent constants, as is the case for isotropic materials. Defected assemblies (B and C) are also isotropic because the defects are chosen so that the removed contacts will not disrupt the symmetry in intergranular forces.

For each assembly, a unit cell is defined. By periodically repeating the unit cell in both horizontal and vertical directions, the complete assembly is formed. The unit cells can be seen in Figure 3 by shaded grains. Discrete simulations of these unit cells with periodic boundary conditions have been performed. In these simulations, intergranular stiffness coefficients are assigned as $\overline{K}_n = 2\overline{K}_s = 17.5$ KN/mm and $\overline{K}_{ns} = \overline{G} = 0.0$. For the three assemblies, multiple simulations, as discussed in Section 3.2, have been performed, and stiffness tensors corresponding to average displacement gradient, fluctuations in displacement gradient, and second gradients are derived and given in Table 1. For assembly A, since all grains in the unit cell are in fact boundary grains, stiffness tensors corresponding to displacement fluctuations and second gradient terms are zero. Indeed, in such grain assembly, all grains always move according to the average displacement gradient applied on the boundary grains. For the two other assemblies, however, not all grains inside the unit cell are boundary grains. Thus, effects of grain displacement fluctuations and second gradient of displacement will be nonzero. Components of the fourth-rank stiffness tensors, $\boldsymbol{C}^M$ and $\boldsymbol{C}^m$, and the sixth-rank stiffness tensor corresponding to second gradient terms, $\boldsymbol{A} = \boldsymbol{A}^u + \boldsymbol{A}^g$, for these regular assemblies are presented in Table 1. It is noteworthy that the stiffness components associated with fluctuation and second gradient increase as more defects are introduced into the assembly.

C_{ijkl}^M and C_{ijkl}^m				A_{ijklmn}		
Component	Assembly A	Assembly B	Assembly C	Component	Assembly B	Assembly C
C_{11}^M	26.5	19.9	17.7	A_{111111}	8.2	60.4
C_{22}^M	26.5	19.9	17.7	A_{112112}	24.7	14.6
C_{33}^M	11.4	8.5	7.6	A_{122122}	74.0	43.9
C_{12}^M	3.8	2.8	2.5	A_{211211}	11.5	56.5
C_{13}^M	0	0.0	0.0	A_{212212}	34.5	20.5
C_{23}^M	0	0.0	0.0	A_{222222}	103.6	61.4
C_{11}^m	0	1.6	7.6	A_{111112}	14.2	4.2
C_{22}^m	0	6.6	8.8	A_{111122}	24.7	7.3
C_{33}^m	0	4.7	6.3	A_{111211}	0.0	1.7
C_{44}^m	0	2.2	7.6	$A_{111212} = A_{112211}$	0.0	−1.5
C_{12}^m	0	0.0	−0.3	$A_{111222} = A_{122211}$	0.0	−2.5
C_{13}^m	0	2.7	2.7	A_{112122}	42.7	25.3
C_{14}^m	0	0.0	0.0	A_{211212}	19.9	5.9
C_{23}^m	0	0.0	0.0	A_{211222}	34.5	10.2
C_{24}^m	0	3.8	3.8	$A_{212222} = A_{221222}$	59.8	59.8
C_{34}^m	0	0.0	−0.3	otherwise	0.0	0.0

Table 1. Components of two fourth-rank stiffness tensors ($\boldsymbol{C}^M$ and $\boldsymbol{C}^m$), in units of GPa, and the sixth-rank second gradient stiffness tensor ($A = A^g + A^u$), in units of N, for the regular unit cells.

4.2. *Irregular tridisperse grain assembly.*

Three random assemblies of grains of three different sizes have been made and analyzed. The three assemblies are composed of 288, 576, and 1152 grains, as shown in Figure 4. In all three assemblies, 52% of grains have a radius of 19 μm, 26% have a radius of 22 μm, and 22% have
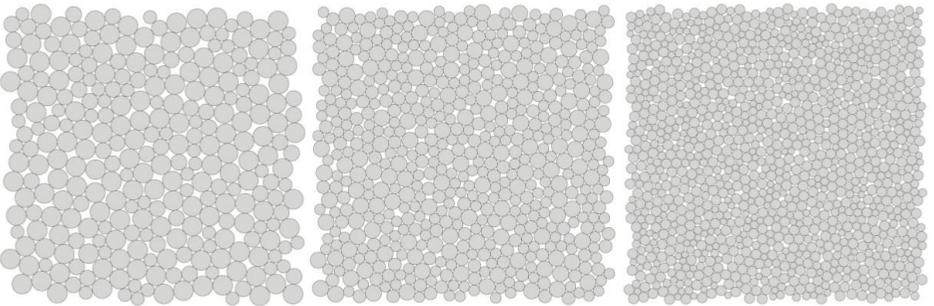

Figure 4. Irregular grain assemblies with different numbers of grains, with constant ratios of grains with different sizes.

component	N = 288			N = 576			N = 1152		
	unit cell	RVE	model results	unit cell	RVE	model results	unit cell	RVE	model results
C_{11}^M	17.9	17.7	17.7	18.9	19.0	19.0	18.1	18.3	18.3
C_{22}^M	17.8	17.7	17.7	19.0	19.0	19.0	18.4	18.3	18.3
C_{33}^M	7.4	7.4	7.4	7.9	7.9	7.9	7.6	7.6	7.6
C_{12}^M	2.8	2.8	2.8	3.1	3.1	3.1	3.1	3.0	3.0
C_{13}^M	0.3	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0
C_{23}^M	0.0	0.0	0.0	0.1	0.0	0.0	−0.1	0.0	0.0
C_{11}^m	4.0	4.2	4.2	3.1	2.9	2.9	2.0	2.0	2.0
C_{22}^m	4.5	4.2	4.2	2.8	2.9	2.9	2.0	2.0	2.0
C_{33}^m	3.8	3.7	3.7	2.4	2.6	2.6	1.7	1.8	1.8
C_{44}^m	3.5	3.7	3.7	2.7	2.6	2.6	1.7	1.8	1.8
C_{12}^m	0.2	0.2	2.8	0.1	0.2	3.1	0.1	0.1	3.0
C_{13}^m	−0.2	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
C_{14}^m	−0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C_{23}^m	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C_{24}^m	−0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C_{34}^m	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1

Table 2. Components of two fourth-rank stiffness tensors (C^M and C^m) for the irregular grain assemblies. All results have units of GPa. The unit cells are as given in Figure 4, the RVE is taken to be composed of randomly oriented unit cells, as in isotropic polycrystals (note the unit cells are amorphous with periodic boundaries), and model results pertain to the RVE.

a radius of 28 μm. Intergranular stiffness coefficients used for discrete simulations are $\overline{K}_n = 2\overline{K}_s = 17.5$ KN/mm and $\overline{K}_{ns} = \overline{G} = 0.0$.

In Tables 2 and 3, the stiffness tensor components for the three assemblies with different numbers of grains are presented. For each assembly, first the components of the unit cell's stiffness tensor are given in the column titled "unit cell". Further, to identify the grain-pair stiffness relevant to the isotropic enhanced continuum model, we consider an RVE of granular media which comprises randomly oriented unit cells, as in polycrystals. The isotropic stiffness tensor of this RVE is estimated through the Voigt–Reuss–Hill (VRH) directional averaging process [Hill 1952]. For each assembly, in the second column in Tables 2 and 3, the components of stiffness tensors of the resulting RVE are presented. It is noteworthy that the "unit cell" results and the estimated RVEs have small differences. Finally, for each assembly, in the last columns, titled "Model results", we give the stiffness tensors

component	N = 288			N = 576			N = 1152		
	unit cell	RVE	model results	unit cell	RVE	model results	unit cell	RVE	model results
A_{111111}	206.9	257.9	257.9	97.1	100.2	100.2	33.3	33.9	33.9
A_{112112}	146.6	130.5	81.8	53.6	48.2	31.7	19.3	16.9	10.6
A_{122122}	247.8	233.1	233.1	82.7	90.3	90.3	25.7	29.9	29.9
A_{211211}	182.4	233.1	233.1	84.9	90.3	90.3	28.1	29.9	29.9
A_{212212}	143.2	130.5	81.8	52.6	48.2	31.7	19.5	16.9	10.6
A_{222222}	287.3	257.9	257.9	96.8	100.2	100.2	30.6	33.9	33.9
A_{111112}	−19.4	0.0	0.0	−3.7	0.5	0.5	−1.2	−0.1	−0.1
A_{111122}	0.0	−15.4	81.8	3.0	−1.2	31.7	0.3	−1.9	10.6
A_{111211}	−3.3	0.0	0.0	−0.9	−0.9	−0.9	0.1	0.1	0.1
A_{111212}	4.4	6.5	6.2	2.1	2.7	2.5	0.7	1.1	1.0
A_{111222}	−0.8	0.0	0.0	0.8	0.0	0.0	0.2	0.0	0.0
A_{112122}	−4.1	0.0	0.0	6.4	0.5	0.5	1.4	−0.1	−0.1
A_{112211}	3.6	5.9	6.2	1.8	2.2	2.5	0.7	0.9	1.0
A_{112212}	−2.5	0.0	0.0	−0.4	0.0	0.0	0.0	0.0	0.0
A_{112222}	4.9	6.5	6.2	1.7	2.7	2.5	0.8	1.1	1.0
A_{122211}	0.1	0.0	0.0	0.8	0.0	0.0	0.2	0.0	0.0
A_{122212}	4.6	5.9	6.2	1.0	2.2	2.5	0.7	0.9	1.0
A_{122222}	1.0	0.0	0.0	0.6	0.9	0.9	−0.4	−0.1	−0.1
A_{211212}	−22.8	0.0	0.0	−4.8	−0.5	−0.5	−1.2	0.1	0.1
A_{211222}	−1.9	−15.4	81.8	4.2	−1.2	31.7	0.8	−1.9	10.6
A_{212222}	3.7	0.0	0.0	5.4	−0.5	−0.5	1.4	0.1	0.1
A_{221222}	3.7	0.0	0.0	5.4	−0.5	−0.5	1.4	0.1	0.1

Table 3. Components of the sixth-rank second gradient stiffness tensor ($A = A^g + A^u$) for the irregular grain assemblies. All results have units of 10^{-5} KN. The unit cells are as given in Figure 4, the RVE is taken to be composed of randomly oriented unit cells as in isotropic polycrystals (note the unit cells are amorphous with periodic boundaries), and model results pertain to the RVE.

derived from calculated microscale stiffness (i.e., grain-pair stiffness coefficients corresponding to macroscale displacement gradient, fluctuations in displacement gradients, and second gradients of displacement fluctuations). The stiffness tensors for the macroscale displacement gradients and the fluctuations in displacement gradient show a very good agreement. For the second gradient stiffness tensor, the microscale stiffness constants are found by an optimization procedure since the number of independent constants is less than the number of independent components

in the second gradient stiffness tensors. A reasonable agreement is also found for the second gradient constitutive coefficients, and, notably, the orders of the major nonzero terms are in concurrence. We further note the need for nonzero coupling terms, K_{ns}, in the microscale constitutive relationships for replicating the fluctuation and second gradient macroscopic stiffness tensors. Though the effects of these stiffness components are of smaller order than those of the diagonal stiffness coefficients, K_n and K_s, nevertheless they conform to the relationships derived from the present model (see Appendix I). We also note that the microscale moment stiffness coefficients, G, are negligible (of the order 10^{-16} N.mm), which implies that for the particular simulations the grain rotations have an insignificant role. However, this is likely a result of assumed zero rotational stiffness in the discrete simulations.

4.3. *Microscale constants for enhanced continuum model.* Finally, we compare the back-calculated microscale constitutive coefficients in Table 4 with those assigned in the discrete simulations. For the case of regular assembly, we observe that the microscale stiffness constants, K_n^M and K_s^M, relevant to the macrostrain stiffness tensor for assembly A are 13% smaller than those used in discrete simulations. The grain-pair stiffness relevant to the continuum model is clearly influenced by the strongly discrete hexagonal microstructure of the grain assembly. As the regular assembly becomes more defective (B and C), the constants, K_n^M and K_s^M, become smaller. However, notably, the ratio K_s^M/K_n^M remains 2, which is the same as that specified in discrete simulations, thus preserving the Poisson's effect. Moreover, the microscale stiffness constants corresponding to the fluctuation and second gradient stiffness tensors appear and increase as more defects are introduced. The ratio K_s^m/K_n^m no longer remains the same as that specified in discrete simulations, implying a different Poisson's effect associated with the fluctuation behavior. In addition, a coupling of the normal and shear behavior is also revealed. It is clear that the needed microscale constitutive coefficients are significantly affected by heterogeneity introduced by microstructure, so that the continuum modeling with

Assembly	K_n^M	K_s^M	K_n^m	K_s^m	K_{ns}^m	K_n^g	K_s^g	K_{ns}^g
A	15.2	7.6						
B	10.9	5.4	2.0	1.7	2.6×10^{-1}	0.3	0.3	4.1×10^{-2}
C	8.4	4.2	3.3	3.0	2.3×10^{-1}	0.4	0.4	-3.6×10^{-2}
$N = 288$	16.8	7.5	3.6	2.8	-4.9×10^{-3}	0.8	0.7	1.2×10^{-4}
$N = 576$	17.1	7.5	2.3	1.9	-2.6×10^{-2}	0.6	0.5	-7.7×10^{-3}
$N = 1152$	16.8	7.2	1.7	1.3	1.1×10^{-2}	0.4	0.3	2.5×10^{-3}

Table 4. Microscale constitutive coefficients derived for both the regular and irregular assemblies with $\bar{K}_n = 2\bar{K}_s = 17.5$ KN/mm and $\bar{K}_{ns} = \bar{G} = 0.0$ (K in kN/mm and G in N.mm).

only macroscale displacement gradients is not sufficient even for the relatively uniform system analyzed here, wherein all grain pairs are given the same stiffness constants. Similar observations can be made from the comparison of microscale stiffness constants of the three irregular assemblies in which smaller particle numbers imply greater microstructural heterogeneity. Clearly, the grain neighborhoods affect the behavior of grain interactions significantly, so that the effective grain-pair stiffness cannot be just estimated from the stiffness of two isolated grains. Additional macroscale deformation measures are, therefore, necessary for the continuum modeling of granular materials. Similar observations have been made for other material systems such as pantographic trusses [Alibert et al. 2003; Seppecher et al. 2011], biomaterials [Andreaus et al. 2012; 2015a], and in fiber composites [Ferretti et al. 2014]. Further, we note that the application of isolated grain-pair potentials or stiffness functions for estimating the energies and stresses associated with different deformation measures, as proposed in some multiscale models that aim to bridge discrete-continuum models, need to be reconsidered in the light of the findings of this paper.

5. Summary and conclusion

The granular micromechanics approach has been used to develop an enhanced continuum model of grain assemblies by the identification of grain (microscale) motions in terms of the macroscale displacement gradient, the fluctuations in displacement gradient as well as their second gradient. Thus, additional stress tensors conjugate to the strain measures as well as additional force measures that are conjugate to the grain-pair displacements are introduced. The expressions of stress tensors are found in terms of the corresponding force and geometric measures, which are different from those derived using the generalized virial theorem. Further, defining the macroscopic strain energy density as the volume average of grain-pair energy functions, macroscopic stiffness tensors corresponding to the kinematic measures have been derived. The continuum stiffness tensors are obtained in terms of grain-pair stiffness coefficients and fabric parameters defining the geometry of grains and their contacts.

To identify the elastic constants of the enhanced continuum model, we perform numerical experiments on grain assemblies using discrete simulations subjected to relevant boundary conditions. The need for additional macroscale deformation measures for the continuum modeling of granular materials becomes evident in this identification process. The obtained elastic constants are then used to determine the microscale (or grain-pair) stiffness coefficients applicable to the continuum model. These grain-scale stiffness coefficients are found to be affected by the heterogeneity of microstructure, as shown by the results for regular grain assembly

made increasingly heterogeneous by introducing vacancy-type defects. The grain-pair stiffness coefficients are clearly different from the isolated grain-pair stiffness used in discrete simulation. These effective stiffness coefficients are unique for each material or granular assembly, since the intergranular mechanisms are affected not only by the two grains under consideration, but also by the grains in the neighborhood, and, by extension, the whole assembly. We also find that the stiffness coefficients corresponding to the average displacement gradient terms have the largest values but the coefficients corresponding to displacement gradient fluctuations are also significant and of a similar order.

Finally, we note that the micromorphic and second gradient terms are necessary for modeling some frequency-dependent wave transmission/reflection phenomena at material interfaces [Misra and Poorsolhjouy 2015b; Placidi et al. 2014; dell'Isola et al. 2012; Madeo et al. 2015]. The micromorphic behavior of granular materials indicates the possibility of realizing materials with alternate synthesis pathways which show specific wave propagation behaviors that can be used for vibration control as an alternative to piezoelectric materials [Maurini et al. 2004; 2006; Porfiri et al. 2005; Vidoli and dell'Isola 2001; Madeo et al. 2014; dell'Isola and Vidoli 1998; Greco et al. 2014] or for damage identification [Ferretti et al. 2014; Andreaus and Baragatti 2011; 2012]. Alternatively, such materials can be applied to help optimize control procedures [Andreaus et al. 2012; 2015a] or for optimal biomaterial design in bone mechanics [Andreaus et al. 2015b]. The identification process described herein can be used for extending the applicability of micromorphic models or their micropolar and second gradient simplifications to describe post-instability macroscale behavior, such as boundary and localization layers in microstructured media [Altenbach et al. 2010; Placidi 2015; Yang et al. 2011; Yang and Misra 2012].

Appendix I. Expressions for elastic constants of 2D granular assemblies

For defining grain-scale constitutive equations, a local coordinate system is introduced for each grain-pair interaction. This coordinate system is composed of a unit normal vector, n_i, in the direction of the branch vector joining the two grains' centroids and another unit vector, s_i, lying in the direction of the tangential plane, whose normal vector is $\boldsymbol{n}$. Cartesian components of the unit vectors are defined as

$$n_i = \langle \cos\theta, \sin\theta \rangle, \quad s_i = \langle -\sin\theta, \cos\theta \rangle. \tag{28}$$

Following the previous discussion about using an average value for the geometric tensors l_i and J_{ij} in any given orientation, and using the 2D density distribution function introduced in (17a), the summations in equations (14)–(16) are rewritten

as integrals in the forms

$$C_{ijkl}^{M} = \frac{1}{V} \sum_{\alpha} K_{ik}^{M} l_{l}^{\alpha} l_{j}^{\alpha} = l^{2} N_{p} \int_{\theta=0}^{2\pi} (K_{ik}^{M} n_{j} n_{l}) \xi \, d\theta, \tag{29}$$

$$C_{ijkl}^{m} = \frac{1}{V} \sum_{\alpha} K_{ik}^{m} l_{l}^{\alpha} l_{j}^{\alpha} = l^{2} N_{p} \int_{\theta=0}^{2\pi} (K_{ik}^{m} n_{j} n_{l}) \xi \, d\theta, \tag{30}$$

and

$$A_{ijklmn}^{g} = \frac{1}{V} \sum_{\alpha} K_{il}^{g} J_{mn}^{\alpha} J_{jk}^{\alpha} = \frac{l^{4} N_{p}}{4} \int_{\theta=0}^{2\pi} (K_{il}^{g} n_{j} n_{k} n_{m} n_{n}) \xi \, d\theta, \tag{31a}$$

$$A_{ijklmn}^{u} = \frac{1}{V} \sum_{\alpha} G_{pq}^{u} e_{mlq} e_{jip} l_{k} l_{n} = l^{2} N_{p} \int_{\theta=0}^{2\pi} (G_{pq}^{u} e_{lmq} e_{ijp} n_{k} n_{n}) \xi \, d\theta. \tag{31b}$$

Note that, since the method is applied here for only 2D modeling, the indices i, j, k, l, m, and n take the value of either 1 or 2. It should be noted here that in a 2D domain (with in-plane coordinate axes 1 and 2) the only possible rotation is the rotation about the axis normal to the plane under consideration, 3 axis, denoted as θ_3, which gives rise to the moment component m_3. Now, in (31b), e_{ijp} and e_{lmq} denote permutation symbols, and since the indices i, j, l, and m can take only values of 1 and 2, the indices p and q should only take the value 3. So the rotational stiffness tensor, G, can have only one component, $G_{pq} = G_{33} = G$. Having this in mind and considering Equation (13) defining the general constitutive equations in intergranular scale, the moment-rotation constitutive equation in grain-scale can thus be written simply as

$$m_3 = G\theta_3. \tag{32}$$

For the force-displacement constitutive equations in microscale, intergranular force and displacement vectors are decomposed in the local directions into two components, one normal and one tangential. Constitutive equations in the local level are then defined in this coordinate system as

$$\begin{Bmatrix} f_n \\ f_s \end{Bmatrix} = \begin{pmatrix} K_n & K_{ns} \\ K_{ns} & K_s \end{pmatrix} \begin{Bmatrix} \delta_n \\ \delta_s \end{Bmatrix}, \tag{33}$$

where we have, for convenience, not shown the superscripts. As it is seen in (33), in the model presented here, the interaction between normal and tangential components of intergranular force and displacement vectors is included. In the previous versions of the method of granular micromechanics this coupling term was ignored, and the stiffness tensor was diagonal and was composed of one normal component, K_n, and one tangential component, K_s. It should be pointed out, however, that the stiffness tensor is still kept symmetric. The stiffness tensor used in (33) should then be rotated to the VE coordinate system to result in the stiffness

tensor K_{ij}, which will be used in the grain-scale constitutive equation $f_i = K_{ij}\delta_j$ and also in equations (29), (30), and (31a), and which is given by

$$K_{ij} = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} = \begin{pmatrix} n_1 & s_1 \\ n_2 & s_2 \end{pmatrix} \begin{pmatrix} K_n & K_{ns} \\ K_{ns} & K_s \end{pmatrix} \begin{pmatrix} n_1 & n_2 \\ s_1 & s_2 \end{pmatrix}. \tag{34}$$

For isotropic materials and in the linear elastic limit, constitutive equations and their corresponding stiffness tensors can be derived in closed form by performing the integrations presented in (29)–(31). The 2D form of the constitutive equations and corresponding stiffness tensors are thus derived as

$$\begin{Bmatrix} \tau_{11} \\ \tau_{22} \\ \tau_{12} \end{Bmatrix} = \begin{pmatrix} C_{11}^M & C_{12}^M & 0 \\ C_{12}^M & C_{11}^M & 0 \\ 0 & 0 & C_{33}^M \end{pmatrix} \begin{Bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{12} \end{Bmatrix}, \qquad \text{where} \quad \begin{cases} C_{11}^M = \frac{1}{8}l^2 N_p(3k_n^M + k_s^M), \\ C_{33}^M = \frac{1}{8}l^2 N_p(k_n^M + k_s^M), \\ C_{12}^M = \frac{1}{8}l^2 N_p(k_n^M - k_s^M), \end{cases} \tag{35}$$

$$\begin{Bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \\ \sigma_{21} \end{Bmatrix} = \begin{pmatrix} C_{11}^m & C_{12}^m & C_{13}^m & C_{14}^m \\ C_{12}^m & C_{11}^m & C_{13}^m & C_{14}^m \\ C_{13}^m & C_{13}^m & C_{33}^m & C_{12}^m \\ C_{14}^m & C_{14}^m & C_{12}^m & C_{33}^m \end{pmatrix} \begin{Bmatrix} \gamma_{11} \\ \gamma_{22} \\ \gamma_{12} \\ \gamma_{21} \end{Bmatrix}, \quad \text{where} \quad \begin{cases} C_{11}^m = \frac{1}{8}l^2 N_p(3k_n^m + k_s^m), \\ C_{33}^m = \frac{1}{8}l^2 N_p(k_n^m + 3k_s^m), \\ C_{12}^m = \frac{1}{8}l^2 N_p(k_n^m - k_s^m), \\ C_{13}^m = -C_{14}^m = \frac{1}{4} - l^2 N_p k_{ns}^m, \end{cases} \tag{36}$$

and

$$\begin{Bmatrix} \mu_{111} \\ \mu_{112} \\ \mu_{121} \\ \mu_{122} \\ \mu_{211} \\ \mu_{212} \\ \mu_{221} \\ \mu_{222} \end{Bmatrix} = ([A_{ijklmn}^g]_{8\times8} + [A_{ijklmn}^u]_{8\times8}) \begin{Bmatrix} \phi_{1,11} \\ \phi_{1,12} \\ \phi_{1,21} \\ \phi_{1,22} \\ \phi_{2,11} \\ \phi_{2,12} \\ \phi_{2,21} \\ \phi_{2,22} \end{Bmatrix} = [A_{ijklmn}]_{8\times8} \begin{Bmatrix} \phi_{1,11} \\ \phi_{1,12} \\ \phi_{1,21} \\ \phi_{1,22} \\ \phi_{2,11} \\ \phi_{2,12} \\ \phi_{2,21} \\ \phi_{2,22} \end{Bmatrix}, \tag{37a}$$

where

$$[A_{ijklmn}^g]_{8\times8} = \frac{1}{16}l^4 N_p \times$$

$$\begin{pmatrix} 5k_n^g + k_s^g & -2k_{ns}^g & -2k_{ns}^g & k_n^g + k_s^g & 4k_{ns}^g & k_n^g - k_s^g & k_n^g - k_s^g & 0 \\ -2k_{ns}^g & k_n^g + k_s^g & k_n^g + k_s^g & -2k_{ns}^g & k_n^g - k_s^g & 0 & 0 & k_n^g - k_s^g \\ -2k_{ns}^g & k_n^g + k_s^g & k_n^g + k_s^g & -2k_{ns}^g & k_n^g - k_s^g & 0 & 0 & k_n^g - k_s^g \\ k_n^g + k_s^g & -2k_{ns}^g & -2k_{ns}^g & k_n^g + 5k_s^g & 0 & k_n^g - k_s^g & k_n^g - k_s^g & -4k_{ns}^g \\ 4k_{ns}^g & k_n^g - k_s^g & k_n^g - k_s^g & 0 & k_n^g + 5k_s^g & -2k_{ns}^g & -2k_{ns}^g & k_n^g + k_s^g \\ k_n^g - k_s^g & 0 & 0 & k_n^g - k_s^g & -2k_{ns}^g & k_n^g + k_s^g & k_n^g + k_s^g & -2k_{ns}^g \\ k_n^g - k_s^g & 0 & 0 & k_n^g - k_s^g & -2k_{ns}^g & k_n^g + k_s^g & k_n^g + k_s^g & -2k_{ns}^g \\ 0 & k_n^g - k_s^g & k_n^g - k_s^g & -4k_{ns}^g & k_n^g + k_s^g & -2k_{ns}^g & -2k_{ns}^g & 5k_n^g + k_s^g \end{pmatrix} \tag{37b}$$

and

$$[A^u_{ijklmn}]_{8\times8} = \frac{l^2 N_p}{2} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & G & 0 & -G & 0 & 0 & 0 \\ 0 & 0 & 0 & G & 0 & -G & 0 & 0 \\ 0 & 0 & -G & 0 & G & 0 & 0 & 0 \\ 0 & 0 & 0 & -G & 0 & G & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{37c}$$

It is noteworthy that the two fourth-rank stiffness tensors, C^M_{ijkl} and C^m_{ijkl}, are formally similar. The only difference between the two tensors is that their components are derived based on grain-pair stiffness coefficients corresponding to different intergranular phenomena. It should also be noted that since the average strain tensor is the symmetric part of the macroscale displacement gradient, the strain and its conjugate stress tensor have three components. Thus, the stiffness matrix linking the average strain tensor to the Cauchy stress, C^M_{ijkl}, is also symmetrized into a 3×3 matrix, while C^m_{ijkl} is a 4×4 matrix.

Appendix II. Variational principle and balance equations

The variation of the internal deformation energy functional in terms of the macroscale measures is obtained as

$$\delta \mathcal{W} = \int_v \delta W \, dV = \int_v (\tau_{ij}\delta\epsilon_{ij} + \sigma_{ij}\delta\gamma_{ij} + \mu_{ijk}\delta\phi_{i,jk}) \, dV$$
$$= -\int_v (\tau_{ij} + \sigma_{ij})_{,j}\delta\bar{\phi}_i \, dV - \int_v (\mu_{ijk,k} + \sigma_{ij})\delta\psi_{ij} \, dV$$
$$+ \int_s (\tau_{ij} + \sigma_{ij})n_j\delta\bar{\phi}_i \, dS + \int_s \mu_{ijk}n_k\delta\psi_{ij} \, dS, \tag{38}$$

where we have used Gauss's divergence theorem and equations (4) and (6). The variation of external energy due to external actions on the system is written as

$$\delta \mathcal{W}^{\text{ext}} = \int_v f_i\delta\bar{\phi}_i \, dV + \int_v \Phi_{ij}\delta\psi_{ij} \, dV + \int_s t_i\delta\bar{\phi}_i \, dS + \int_s T_{ij}\delta\psi_{ij} \, dS, \tag{39}$$

where f_i is the noncontact volumic (body) force per unit volume, t_i is the contact traction, defined as a surface force per unit area, Φ_{ij} is the noncontact volumic (body) double force per unit volume, and T_{ij} is the contact double traction, defined

as double force per unit area. Combining equations (38) and (39) results in

$$\int_v [f_i + (\tau_{ij} + \sigma_{ij})_{,j}]\delta\bar{\phi}_i\, dV + \int_v [\Phi_{i,j} + (\mu_{ijk,k} + \sigma_{ij})]\delta\psi_{ij}\, dV$$

$$+ \int_s [t_i - (\tau_{ij} + \sigma_{ij})n_j]\delta\bar{\phi}_i\, dS + \int_s [T_{ij} - \mu_{ijk}n_k]\delta\psi_{ij}\, dS = 0, \quad (40)$$

which leads to the following balance equations and traction boundary conditions:

$$\begin{cases} (\tau_{ij} + \sigma_{ij})_{,j} + f_i = 0, \\ \mu_{ijk,i} + \sigma_{jk} + \Phi_{jk} = 0, \end{cases} \qquad \begin{cases} (\tau_{ij} + \sigma_{ij})n_j = t_i, \\ \mu_{ijk}n_k = T_{jk}. \end{cases}$$

References

[Alibert et al. 2003] J.-J. Alibert, P. Seppecher, and F. Dell'Isola, "Truss modular beams with deformation energy depending on higher displacement gradients", *Math. Mech. Solids* **8**:1 (2003), 51–73.

[Altenbach et al. 2010] H. Altenbach, V. A. Eremeyev, L. P. Lebedev, and L. A. Rendón, "Acceleration waves and ellipticity in thermoelastic micropolar media", *Arch. Appl. Mech.* **80**:3 (2010), 217–227.

[Andreaus and Baragatti 2011] U. Andreaus and P. Baragatti, "Cracked beam identification by numerically analysing the nonlinear behaviour of the harmonically forced response", *J. Sound Vib.* **330**:4 (2011), 721–742.

[Andreaus and Baragatti 2012] U. Andreaus and P. Baragatti, "Experimental damage detection of cracked beams by using nonlinear characteristics of forced response", *Mech. Syst. Signal Process.* **31** (2012), 382–404.

[Andreaus et al. 2012] U. Andreaus, M. Colloca, and D. Iacoviello, "An optimal control procedure for bone adaptation under mechanical stimulus", *Control Eng. Practice* **20**:6 (2012), 575–583.

[Andreaus et al. 2015a] U. Andreaus, M. Colloca, and D. Iacoviello, "Optimal bone density distributions: numerical analysis of the osteocyte spatial influence in bone remodeling", *Comput. Methods Programs Biomedicine* **113**:1 (09/08 2015), 80–91.

[Andreaus et al. 2015b] U. Andreaus, I. Giorgio, and A. Madeo, "Modeling of the interaction between bone tissue and resorbable biomaterial as linear elastic materials with voids", *Z. Angew. Math. Phys.* **66**:1 (2015), 209–237.

[Auffray et al. 2015] N. Auffray, F. dell'Isola, V. A. Eremeyev, A. Madeo, and G. Rosi, "Analytical continuum mechanics á la Hamilton–Piola least action principle for second gradient continua and capillary fluids", *Math. Mech. Solids* **20**:4 (2015), 375–417.

[Cauchy 1826–1830] A.-L. Cauchy, "Sur l'equilibre et le mouvement d'un systeme de points materiels sollicites par des forces d'attraction ou de repulsion mutuelle", pp. 188–212 in *Excercises de mathemátiques*, vol. 3, 1826–1830.

[Chang and Misra 1989] C. S. Chang and A. Misra, "Computer simulation and modelling of mechanical properties of particulates", *Comput. Geotech.* **7**:4 (1989), 269–287.

[Chen and Lee 2003] Y. Chen and J. D. Lee, "Connecting molecular dynamics to micromorphic theory, (I): Instantaneous and averaged mechanical variables", *Physica A* **322** (2003), 359–376.

[Cosserat and Cosserat 1909] E. Cosserat and F. Cosserat, *Théorie des corps deformables*, A. Hermann, Paris, 1909. Translated as Theory of deformable bodies, Paris, A. Hermann, 1909.

[dell'Isola and Vidoli 1998] F. dell'Isola and S. Vidoli, "Continuum modelling of piezoelectrome-chanical truss beams: an application to vibration damping", *Arch. Appl. Mech.* **68**:1 (1998), 1–19.

[dell'Isola et al. 2012] F. dell'Isola, A. Madeo, and L. Placidi, "Linear plane wave propagation and normal transmission and reflection at discontinuity surfaces in second gradient 3D continua", *Z. Angew. Math. Mech.* **92**:1 (2012), 52–71.

[dell'Isola et al. 2014] F. dell'Isola, U. Andreaus, and L. Placidi, "At the origins and in the vanguard of peridynamics, non-local and higher-gradient continuum mechanics: an underestimated and still topical contribution of Gabrio Piola", *Math. Mech. Solids* **20**:8 (2014), 887–928.

[Eringen 1999] A. C. Eringen, *Microcontinuum field theories, I: Foundations and solids*, Springer, New York, 1999.

[Ferretti et al. 2014] M. Ferretti, A. Madeo, F. dell'Isola, and P. Boisse, "Modeling the onset of shear boundary layers in fibrous composite reinforcements by second-gradient theory", *Z. Angew. Math. Phys.* **65**:3 (2014), 587–612.

[Ganghoffer 2010] J.-F. Ganghoffer, "On the generalized virial theorem and Eshelby tensors", *Int. J. Solids Struct.* **47**:9 (2010), 1209–1220.

[Germain 1973] P. Germain, "The method of virtual power in continuum mechanics, 2: Microstructure", *SIAM J. Appl. Math.* **25**:3 (1973), 556–575.

[Greco et al. 2014] L. Greco, N. Impollonia, and M. Cuomo, "A procedure for the static analysis of cable structures following elastic catenary theory", *Int. J. Solids Struct.* **51**:7–8 (2014), 1521–1533.

[Green and Rivlin 1964] A. E. Green and R. S. Rivlin, "Multipolar continuum mechanics", *Arch. Ration. Mech. An.* **17** (1964), 113–147.

[Hill 1952] R. Hill, "The elastic behaviour of a crystalline aggregate", *P. Phys. Soc. A* **65**:5 (1952), 349.

[Jenkins et al. 2005] J. Jenkins, D. Johnson, L. La Ragione, and H. Makse, "Fluctuations and the effective moduli of an isotropic, random aggregate of identical, frictionless spheres", *J. Mech. Phys. Solids* **53**:1 (2005), 197–225.

[Madeo et al. 2014] A. Madeo, L. Placidi, and G. Rosi, "Towards the design of metamaterials with enhanced damage sensitivity: second gradient porous materials", *Res. Nondestruct. Eval.* **25**:2 (2014), 99–124.

[Madeo et al. 2015] A. Madeo, P. Neff, I. D. Ghiba, L. Placidi, and G. Rosi, "Wave propagation in relaxed micromorphic continua: modeling metamaterials with frequency band-gaps", *Continuum Mech. Therm.* **27**:4–5 (2015), 551–570.

[Maugin 2014] G. A. Maugin, "Some remarks on generalized continuum mechanics", *Math. Mech. Sol.* **20**:3 (2014), 280–291.

[Maurini et al. 2004] C. Maurini, F. dell'Isola, and J. Pouget, "On models of layered piezoelectric beams for passive vibration control", *J. Phys.* (*France*) *IV* **115** (2004), 307–316.

[Maurini et al. 2006] C. Maurini, J. Pouget, and F. dell'Isola, "Extension of the Euler–Bernoulli model of piezoelectric laminates to include 3D effects via a mixed approach", *Comput. Struct.* **84**:22–23 (2006), 1438–1458. Composite Adaptive Structures: Modelling and Simulation.

[Mindlin 1964] R. D. Mindlin, "Micro-structure in linear elasticity", *Arch. Ration. Mech. An.* **16** (1964), 51–78.

[Misra 1998] A. Misra, "Particle kinematics in sheared rod assemblies", pp. 261–266 in *Physics of dry granular media*, edited by H. Herrmann et al., NATO ASI Series **350**, Springer Netherlands, 1998.

[Misra and Chang 1993] A. Misra and C. S. Chang, "Effective elastic moduli of heterogeneous granular solids", *Int. J. Solids Struct.* **30**:18 (1993), 2547–2566.

[Misra and Jiang 1997] A. Misra and H. Jiang, "Measured kinematic fields in the biaxial shear of granular materials", *Comput. Geotech.* **20**:3–4 (1997), 267–285.

[Misra and Poorsolhjouy 2015a] A. Misra and P. Poorsolhjouy, "Elastic behavior of 2D grain packing modeled as micromorphic media based upon granular micromechanics", *J. Eng. Mech.-ASCE* (2015). (in review).

[Misra and Poorsolhjouy 2015b] A. Misra and P. Poorsolhjouy, "Granular micromechanics based micromorphic model predicts frequency band gaps", *Continuum Mech. Therm.* (2015), 1–20.

[Misra and Poorsolhjouy 2015c] A. Misra and P. Poorsolhjouy, "Granular micromechanics model for damage and plasticity of cementitious materials based upon thermomechanics", *Math. Mech. Solids* (2015).

[Misra and Singh 2014] A. Misra and V. Singh, "Nonlinear granular micromechanics model for multi-axial rate-dependent behavior", *Int. J. Solids Struct.* **51**:13 (2014), 2272–2282.

[Misra and Singh 2015] A. Misra and V. Singh, "Thermomechanics-based nonlinear rate-dependent coupled damage-plasticity granular micromechanics model", *Continuum Mech. Therm.* **27**:4–5 (2015), 787–817.

[Navier 1827] C. L. Navier, "Sur les lois de l'equilibre et du mouvement des corps solides elastiques", pp. 375–393 in *Memoire de l'academie royale de sciences*, vol. 7, 1827.

[Peters and Walizer 2013] J. F. Peters and L. E. Walizer, "Patterned nonaffine motion in granular media", *J. Eng. Mech.* (*ASCE*) **139**:10 (2013), 1479–1490.

[Placidi 2015] L. Placidi, "A variational approach for a nonlinear 1-dimensional second gradient continuum damage model", *Continuum Mech. Therm.* **27**:4–5 (2015), 623–638.

[Placidi et al. 2014] L. Placidi, G. Rosi, I. Giorgio, and A. Madeo, "Reflection and transmission of plane waves at surfaces carrying material properties and embedded in second-gradient materials", *Math. Mech. Solids* **19**:5 (2014), 555–578.

[Porfiri et al. 2005] M. Porfiri, F. dell'Isola, and E. Santini, "Modeling and design of passive electric networks interconnecting piezoelectric transducers for distributed vibration control", *Int. J. Appl. Electrom.* **21** (2005), 69–87.

[Seppecher et al. 2011] P. Seppecher, J.-J. Alibert, and F. dell'Isola, "Linear elastic trusses leading to continua with exotic mechanical interactions", *J. Phys. Conf. Ser.* **319**:1 (2011), 012018.

[Toupin 1964] R. A. Toupin, "Theories of elasticity with couple-stress", *Arch. Ration. Mech. An.* **17** (1964), 85–112.

[Trentadue 2001] F. Trentadue, "A micromechanical model for a non-linear elastic granular material based on local equilibrium conditions", *Int. J. Solids Struct.* **38**:40–41 (2001), 7319–7342.

[Vidoli and dell'Isola 2001] S. Vidoli and F. dell'Isola, "Vibration control in plates by uniformly distributed PZT actuators interconnected via electric networks", *Eur. J. Mech. A Solids* **20**:3 (2001), 435–456.

[Yang and Misra 2012] Y. Yang and A. Misra, "Micromechanics based second gradient continuum theory for shear band modeling in cohesive granular materials following damage elasticity", *Int. J. Solids Struct.* **49**:18 (2012), 2500–2514.

[Yang et al. 2011] Y. Yang, W.-Y. Ching, and A. Misra, "Higher-order continuum theory applied to fracture simulation of nanoscale intergranular glassy film", *J. Nanomech. Micromech.* **1**:2 (2011), 60–71.

ANIL MISRA: amisra@ku.edu
Civil, Environmental and Architectural Engineering Department, University of Kansas,
1530 W. 15th Street, Learned Hall, Lawrence, KS 66045-7609, United States

PAYAM POORSOLHJOUY: payam_poorsolhjouy@ku.edu
Civil, Environmental and Architectural Engineering Department, University of Kansas,
1530 W. 15th Street, Learned Hall, Lawrence, KS 66045-7609, United States

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the submission page.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in MEMOCS are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and a Mathematics Subject Classification or a Physics and Astronomy Classification Scheme code for the article, and, for each author, postal address, affiliation (if appropriate), and email address if available. A home-page URL is optional.

Format. Authors are encouraged to use LaTeX and the standard amsart class, but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should normally be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages — Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc. — allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with as many details as you can about how your graphics were generated.

Bundle your figure files into a single archive (using zip, tar, rar or other format of your choice) and upload on the link you been provided at acceptance time. Each figure should be captioned and numbered so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript with all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables.

White Space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

M∩M